

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2005, 05:58:22 ; Search time 327 Seconds
(without alignments)
9157.145 Million cell updates/sec

Title: US-10-068-870A-4
Perfect score: 1830
Sequence: 1 atgaaaaaagaatttagc.....tgaagattattagatg 1830

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.6	5.2	5361	3	US-08-973-462-2
2	95.6	5.2	6152	3	US-08-973-462-1
3	87	4.8	3095	6	5231168-1
4	87	4.8	3095	6	5231168-1
5	85.2	4.7	8155	4	US-08-956-171E-63
6	85.2	4.7	8155	4	US-08-781-986A-63
7	82.2	4.5	2370	4	US-09-754-947-2
8	76	4.2	1039	4	US-09-902-540-1280
9	73.4	4.0	4766	5	PCT-US93-07261-10
10	71.6	3.9	14066	4	US-09-601-138-56
11	70	3.8	3279	3	US-08-446-137B-1
12	69.8	3.8	612	4	US-09-902-540-1357
13	69.4	3.8	1891	3	US-08-973-462-3
14	69.4	3.8	2489	3	US-09-141-047-7
15	69.4	3.8	9636	1	US-08-323-170B-1
16	69.4	3.8	9636	3	US-08-954-441-1
17	68.4	3.7	7218	1	US-08-232-463-14
18	68.4	3.7	18169	4	US-09-949-016-12776
19	68.4	3.7	191569	4	US-09-949-016-15940
20	67.8	3.7	3057	4	US-09-601-138-55
21	65	3.6	3356	3	US-09-379-523-4
22	65	3.6	580073	4	US-08-545-58D-1
23	64.4	3.5	1431	3	US-09-316-083-2
24	64.4	3.5	1431	4	US-09-933-700-2
25	63.6	3.5	1956	3	US-08-559-896B-1
26	63.6	3.5	1956	4	US-09-351-794A-1
27	63	3.4	19124	2	US-08-487-826B-13

C	28	62.8	3.4	3543	4	US-09-710-279-4149	Sequence 4149, Ap
	29	62.8	3.4	3840	4	US-09-710-279-3187	Sequence 3187, Ap
	30	62.8	3.4	11091	3	US-09-134-001C-2243	Sequence 2243, Ap
	31	62.8	3.4	205044	4	US-09-949-016-15851	Sequence 15851, A
	32	62.8	3.4	205044	4	US-09-949-016-15852	Sequence 15852, A
	33	62.8	3.4	205044	4	US-09-949-016-15853	Sequence 15853, A
	34	62.8	3.4	223471	4	US-09-949-016-12387	Sequence 12387, A
	35	62.8	3.4	223471	4	US-09-949-016-12724	Sequence 12724, A
	36	62.8	3.4	223471	4	US-09-949-016-12725	Sequence 12725, A
	37	62.6	3.4	3489	2	US-08-728-323A-1	Sequence 1, Appli
	38	62.6	3.4	3489	3	US-09-298-568-1	Sequence 1, Appli
	39	62.6	3.4	3489	4	US-09-410-399-1	Sequence 1, Appli
	40	62.6	3.4	3489	4	US-09-894-273-1	Sequence 1, Appli
C	41	62.6	3.4	32207	2	US-08-770-379-20	Sequence 20, Appl
C	42	62.6	3.4	32207	3	US-08-757-699A-20	Sequence 20, Appl
C	43	62.6	3.4	32207	3	US-09-230-371A-20	Sequence 31, Appl
C	44	61.4	3.4	810	4	US-09-601-198-31	Sequence 30531, A
C	45	60.8	3.3	601	4	US-09-949-016-30531	

ALIGNMENTS

RESULT 1

US-08-973-462-2
; Sequence 2, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
; US-08-973-462-2

Query Match	5.2%	Score 95.6;	DB 3;	Length 5361;
Best Local Similarity	43.6%	Pred. No. 6e-11;		
Mismatches	645;	Conservative	0;	Mismatches 809; Indels 24; Gaps 4;
Qy	233	CAAAATGAGAAGATCTTCAACTCCAACTGGAGATATAAAAGTTTATTGAGAACAACTT	292	
Db	770	CAAGTGTGGAAGAAAGTATAGCTTCAAGTGTGGAAGTATAGATTTCAAGTATTGAAG	829	
Qy	293	TAACACAGCTAATGGAATGAAGTATATCTAAGACAA---CTTTAAAAATTTAGATG	349	
Db	830	AAATGTAGCTCAACTGTGGAAGAAATCGTAGCTCCAGTGTGTTGAGAAGTGTGCTC	889	
Qy	350	CAGGAGAATATGCTATTATAGATTTAACTTTATATATGCTTAAACTGTTGAAATTAAG	409	
Db	890	CAAGTGTGGAAGAAAGTATAGAGAAATCTTGAAGAAAGTATAGTGAAGTGTGAAG	949	
Qy	410	TAGTAGCAGTCTAGTGAAGAAACAGTAGTGTGTTCTAGTATGCGAAAGTATGTCAGAA	469	
Db	950	AAAGTGTAGCTGAAAATGTTTGAAGAAAGTGTAGCTGAAAATGTTTGAAGAAAGTGTAG	1009	
Qy	470	ATATAGCTGAAAATATGTTTGAAGAAAGTATAGAAAGTATAGAAATGCACTAAACTATA	529	
Db	1010	AAATGTGGAAGAAATCGTAGCTCCAACTGTTTGAAGAAATCGTAGCTCCAACTGTTGA	1069	

QY 530 ATGCTCAGATTTTCAGTAAACCTGATAGTTACTATCAAGTAGTCTTTTATCCAAAGGAA 589
 Db 1070 AAATTGTAGCTCCAAAGTGTGTGAAGAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAG 1129
 QY 590 AGAGATTACAAGGTTTCTCAACTTTATAGAGCTACAAATATATATGAAGAACTGCAATATG 649
 Db 1130 AAAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAATAATGTTGAAG 1189
 QY 650 GTAATACACAGTAATATTAACCTCTAAATCTACTAGTAAGAGTAATTTAAAGACTGCAG 709
 Db 1190 AAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTG 1249
 QY 710 TAGAAGAGTTACAAAATTTGAAGCTAGTTATCTTAATACTACAACTTTAGCTGGTGTAG 769
 Db 1250 AAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1309
 QY 770 ACAGAAATACAAACAGCTATAGAGATAAGTAAGAAATATTACAAATATGATGGCGAGAAAT 829
 Db 1310 AAAATGTAGCTCCAAAGTGTGTGAAGAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAG 1369
 QY 830 CAGATCAATCAGCTGATGTTAAAGAGAAATGTTAAAGAAATGTTGATTTAGTAGTGCAATG 889
 Db 1370 AAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1429
 QY 890 CACTAGTAGATGATTTAGTTGGCTCCTTTAGCAGCAGAAAGATGCTCCACTATTAT 949
 Db 1430 AAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTG 1489
 QY 950 TAACTTCAAAAGATAAATTTAGATTTCTGTCAGTAAA-----ATCTGAAATAAAGAGAG 1000
 Db 1490 AAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1549
 QY 1001 TTTTAGACTTAAACCTTCAACAGAAAGTACAGGAAAAACAGTTTATATAGCTGGTGGAG 1060
 Db 1550 AAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1609
 QY 1061 TTAATAGTCTATCTAAGAAAGTGTGAAGAAATGAGAAATCAATGGATTTAAAGTTGAA 1120
 Db 1610 CAAGTGTGTGAAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAG 1669
 QY 1121 GATTTCTCAGTGATGATGATGAAATCTTTTAAATAATAGCAGGTGAAATAGGCTTAG 1180
 Db 1670 AAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTG 1729
 QY 1181 ATAATGATAGGCTTATAGTTGTGGGAAACAGGAATAGCAGATGCAATGATATAGCTT 1240
 Db 1730 AAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1789
 QY 1241 CAGTTGCTTCTACTAAATTTAGATGCTAATG-----GTGTTGTAGATAGAACAAATG 1291
 Db 1790 AAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAGTGTGTGAAGAAAGTGTGGCTC 1849
 QY 1292 GACATGCTACTCCAAATGTTGTTAGATGGAAGAAAGCTGATAAATAATCTGATGACTTAG 1351
 Db 1850 CAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1909
 QY 1352 ATAGTTTCTTAGAAGCGCTGATGTAGATATATAGTGGATTTGCAAGTGTATCTGAAA 1411
 Db 1910 AAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1969
 QY 1412 AGATGGAAGAGCTATATCAGATGCTGCTGTAAGGCGTTACAGAGTTTAAAGCGGAG 1471
 Db 1970 CAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAGTGTGTAG 2029
 QY 1472 ATAGACACAGACACTAATCTGAAGTTTATAAACAATATTATGCTAATGATCTGAAAATAG 1531
 Db 2030 AAAAGTGTGCTCAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTGAAG 2086
 QY 1532 CTAAGCTGAGTTTATAGATAAGATTTAGGCTTCAAGTGTAGTGAAGTGAAGAGTATTTA 1591
 Db 2087 CTGAAAATTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTG 2146
 QY 1592 ATTTCCTATGTAGCTAAAGATGGATCTACAAAGAGATCAATTTAGTTGATGCAATTAGCAG 1651

Db 2147 AAGAAAGTGTAGCTGAAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAG 2206
 QY 1652 TAGGAGCTGTTCTCGATATATAAACTTGTCTCCAGTTGTA 1689
 Db 2207 CTCCTCACTGTTGAAGAAATGTTAGCTCCAAAGTGTGTA 2244

RESULT 2
 US-08-973-462-1
 ; Sequence 1, Application US/08973462B
 ; Patent No. 6151270
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUILHE, PIERRE
 ; APPLICANT: DAUBERSIES, PIERRE
 ; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
 ; FILE REFERENCE: 0660-0125-0 PCT
 ; CURRENT APPLICATION NUMBER: US/08/973,462B
 ; CURRENT FILING DATE: 1998-02-06
 ; EARLIER APPLICATION NUMBER: PCT/FR96/00894
 ; EARLIER FILING DATE: 1996-06-12
 ; EARLIER APPLICATION NUMBER: FR 95/07007
 ; EARLIER FILING DATE: 1995-06-13
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 6152
 ; TYPE: DNA
 ; ORGANISM: P. falciparum
 ; US-08-973-462-1

Query Match 5.2%; Score 95.6; DB 3; Length 6152;
 Best Local Similarity 43.6%; Pred. No. 6.2e-11;
 Matches 645; Conservative 0; Mismatches 809; Indels 24; Gaps 4;

QY 233 CAAATGAGAAAGATTTCTTCACTCCAACTGGAGATAAAAAAGTTTATTTCAGAACAAACTT 292
 Db 1014 CAAGTGTGAAGAAAGTAGCTTCAAGTGTGATGAAGATATAGATTCAAGTATTGAAG 1073
 QY 293 TAACTACAGCTAAATGAAATGAAGATTTATGAAGACAA---CTTTAAAAATTTAGATG 349
 Db 1074 AAAATGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAAAGTGTGTAGAAAGTGGCTC 1133
 QY 350 CAGGAGAAATATGCTATATATAGATTTAACTATATAATAATGCTTAAACTGTTGAAATTAAG 409
 Db 1134 CAAGTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1193
 QY 410 TAGTAGCAGCTAGTCAAAAAACAGTAGTTGTATCTAGTATCGCAAAAAATAGTCGAAAAG 469
 Db 1194 AAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTG 1253
 QY 470 ATATAGCTGAAAAATPATGTTTGAAGACAAAGACTTAGAAAAATGCACCTAAAAACTATAA 529
 Db 1254 AAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1313
 QY 530 ATGCTCTCAGATTTCAAGTAAAACTGATGTTACTATCAAGTAGTTCTTTTATCCAAAGGAA 589
 Db 1314 AAAATGTAGCTCCAAAGTGTGTAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTAGAG 1373
 QY 590 AGAGATTACAAGTTTCTCACTTATAGAGCTACAAAAATTATAATGAAGAACTGCATATG 649
 Db 1374 AAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAG 1433
 QY 650 GTAATACACAGTAATATTAACCTCTAAAACTCTACTAGTAAGAGTAAATTTAAAGACTGCGAG 709
 Db 1434 AAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTG 1493
 QY 710 TAGAAGAGTTACAAAATTTGAATGCTAGTTATTCTTAATCTACAATCTTTAGCTGGTGTAG 769
 Db 1494 AAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1553
 QY 770 ACAGAAATCAAAACAGCTATAGAGATAAGTAAGAAATATTATCAATATGATGCGGAGAAAT 829

1554 AATCTAGCTCAAGTGTGTAGAAAGTGGCTCCAAAGTGTGAAGAAAGTGTAGAAG 1613
QY
830 CAGATCAITTCAGCTGATGTTAAAGAGAAATGTTAAATGTTTATAGTAGTGCAGATG 889
Db
1614 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAG 1673
QY
890 CACTAGTGTAGTGTAGTGTGGCTCTTTAGCAGCAGAGAAAGATGCTCCACTATTAT 949
Db
1674 AAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTG 1733
QY
950 TAACTTTCAAAAGATATAATAGATTTCTGCTAGTAAA-----ATCTGAAATAAGAGAG 1000
Db
1734 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAG 1793
QY
1001 TTTTAGACTTTAAAACTTCAAGAGATTAACAGGAAACAGTTTATATAGCTGGTGGAG 1060
Db
1794 AAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTC 1853
QY
1061 TTAATAGTGTATCTAAGAAAGTGTAAACAGAAATTAGAAATCAATGGGATTTAAAGTTGAAA 1120
Db
1854 CAAAGTGTGTAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAG 1913
QY
1121 GATTTCTCAGGTGATGATAGATATAAGAACTTTTAAATAAGAGTGAAGTAAAGGCTTAG 1180
Db
1914 AAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTG 1973
QY
1181 ATAATGATAAGGCTTATGTAGTGTGGTGGAAACAGATTAACAGATGCTCAGTATAGCTT 1240
Db
1974 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 2033
QY
1241 CAGTGTCTTACTTAATAGATGTTGTAATG-----GTGTTGTAGATAGAAACAAATG 1291
Db
2034 AAATCGTAGCTCAACTGTTGAAGAAATGTAGCTCCAACTGTTGTAGAAAGTGTGGCTC 2093
QY
1292 GACATGCTACTCCAAATAGTGTGTAGATGGAAGAAAGCTGTAAATAATCTGTAGTACTAG 1351
Db
2094 CAAAGTGTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAG 2153
QY
1352 ATAGTGTCTTGAAGAGCGCTGATGATATATATAGTGTGATTTGCAAGTGTATCTGAAA 1411
Db
2154 AAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTC 2213
QY
1412 AGATGAAGAGCTATATCAGATGCTACTGTGAAGCGCTTACAAAGTGTAAAGGCGAGC 1471
Db
2214 CAACTGTGGAAGAAATCGTAGCTCAACTGTGGAAGAAATGTAGCTCCAACTGTTGTAG 2273
QY
1472 ATAGCAAGACACTAACTCTGAAGTATATAAAACATATATGCTTAATGATGATCTGAATAG 1531
Db
2274 AAAGTGTGGCTCCAACTGTTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTG---TAG 2330
QY
1532 CTAAGCTGCGATTTTATAGATAAGATTAAGTGTGCTTCAAGTGTAGTGTGAGGAGTATTTA 1591
Db
2331 CTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTG 2390
QY
1592 ATTTCTATGATCTAAAGATGATCTACAAAGAGATCAATTAAGTGTGATGATCTAGCAG 1651
Db
2391 AAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGGAAGAAATCGTAG 2450
QY
1652 TAGCAGCTGTGGCTGGATATAAACTGTCTCCAGTTGTA 1689
Db
2451 CTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGTA 2488

RESULT 3

5231168-1

; Patent No. 5231168

; APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;

; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.

; TITLE OF INVENTION: MALARIA ANTIGEN

; NUMBER OF SEQUENCES: 19

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/409,658

; FILING DATE: 18-SEP-1989

; SEQ ID NO: 1:
; LENGTH: 3095

5231168-1

Query Match 4.8%; Score 87; DB 6; Length 3095;

Best Local Similarity 43.7%; Pred. No. 3.5e-09;

Matches 436; Conservative 0; Mismatches 555; Indels 1;

QY 165 AGCAGCACAAAATCTTAAACAAAGGTGCAGTTATTTTGAACAAAGATACAAAAAGTTACTTTT 224
Db 1194 AGTTGAACATGAATAGTAGAAGTTGAAGAAATCTACCAGAGATATAAAATGAAAGG 1253
QY 225 CTATGATGCAATGATGAAGATTTCTCAACTCCACTGCGAGATATAAAAGTTTATTCAGA 284
Db 1254 TCAACATGAAATAGTAGAGTTGAAGAAATCTACCAGAGATGATAAAATGAAAGAT 1313
QY 285 ACAACTTTTAACCTACAGCTTAATGGAATGAAGATTTATGTAAAGACAACTTTAAAAAATTT 344
Db 1314 TGAACATGAAATAGTAGAAGTTGAAGAAATCTACCAGAGATATAAAATGAAAGGTC 1373
QY 345 AGATGCGAGAGATATGCTTATTTATAGATTTAACTTTATAATAATGCTTAAACTGTTGAAAT 404
Db 1374 ACATGAAATAGTAGAGTTGAAGAAATCTACCAGAGATATAAAATGAAAGTGTGAACA 1433
QY 405 TAAAGTGTAGCAGCTGTAGTGAAGAAACAGTAGTTGTATCTAGTGATGCGGAAAAATAGTGC 464
Db 1434 TGAATAGTAGAAGTTGAAGAAATCTACCAGAGATATAAAATGAAAGGTCACATGA 1493
QY 465 AAAAGATATAGCTGAAAAATATGTGTTTGAAGACAAAGACTTTAGAAAAATGCCTAAAAAC 524
Db 1494 AATAGTAGAGGTTGAAGAAATCTACCAGAGATATAAAATGAAAGTGTCAACATGAAAT 1553
QY 525 TATAATGCTCAGATTTTCAGTAAAACTGATAGTTACTATCAAGTAGTGTCTTTTATCCAAA 584
Db 1554 AGTAGAAGTTGAAGAAATCTACCAGAGATATAAAATGAAAGGTCACATGAAATAGT 1613
QY 585 AGGAAAGAGATTAACAAGTTTCTCAACTTATAGAGCTACAAATATAAATGAAGAACTGC 644
Db 1614 AGAGTTTGAAGAAATCTACCAGAGATATAAAATGAAAGGTCACATGAAATAGT 1673
QY 645 ATATGTATATACACAGTAAATATTAATCTTAAATCTACTAGTAAGAGTAAATTTAAAGAC 704
Db 1674 AGAGTTTGAAGAAATCTACCAGAGATATAAAATGAAAGGTCACATGAAATAGT 1733
QY 705 TGCAGTAGAGAGCTTACAAAAATTTGAATGCTAGTTTCTAATACTACAACTTTAGCTGG 764
Db 1734 GGTGAAGAAATCTACCAGAGATATAAAATGAAAGGTCACATGAAATAGTAGGT 1793
QY 765 TGATGACAGAAATACAAACAGCTATAGAGATAAGTAAAGAAATATTAATTAATGATGGCGA 824
Db 1794 TGAAGAAATCTTCCAGAAATTTGTTGAATTTGAAGAAAGTACCATCACAAACAAATACAA 1853
QY 825 GAAATCAGATCATTCAGCTGATGTTTAAAGAGAAATGTTAAAAATGTTGTTATAGTAGTGC 884
Db 1854 T-----GAAATATTTGAAACTATAAAACAGAGAAATAAAGAAATGAAATTTAGTGTGA 1907
QY 885 AAATGCACTAGTAGATGATGTTAGTTCGGCTCTCTTGTAGCAGCAGAGAAAGATGCTCCACT 944
Db 1908 AGAAAAAGCAATTTCCACAGAACCCGCTGCTACCTACATTAATGAAATGAAAGCGTTAC 1967
QY 945 ATTATTAACCTTCAAAAGATAAATTAGATTCGTGATGATAAAATCTGAAATAAAGAGATTTT 1004
Db 1968 TCCCAACCACTCTGAAGGTGAATCCACTAAACAGAGATATAGTTCAAAATTAATAAGTACA 2027
QY 1005 AGACTTAAAAATCTTCAACAGAGATAACAGAGAAAAAAGTTTATATAGCTGTTGAGTTAA 1064
Db 2028 AGAAAAATAACCAAAATAAAAGGAAACACCAAGTAGTAGAGTGGTCCAAAAATGTTAGAAC 2087
QY 1065 TAGTGATCTAAGAAAGTTCTTAACAGAAATTAGAATCAATGGGATTTAAAGATTTGAAAGATT 1124
Db 2088 AAATATACAAAGAGATGATTAATGATGAAGAGGATGATGATGATATAGATTTTGAAGATT 2147
QY 1125 CTCAGGTGATGATAGATATGAAACTTCTTTAAAAATA 1161

Db 2148 ATCAAGAAAGATGATGAAAAGGATTCATCAATAAAA 2184

RESULT 4

5231168-1
; Patent No. 5231168
; APPLICANT: DZIGIEL, MORTEN; BORRE, MARTIN; JERSEN, SOREN;
; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/409,658
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 1:
; LENGTH: 3095
5231168-1

Query Match 4.8%; Score 87; DB 6; Length 3095;
Best Local Similarity 43.7%; Pred. No. 3.5e-09;
Matches 436; Conservative 0; Mismatches 555; Indels 6; Gaps 1;

QY 165 AGCAGCACAAAACCTTAACACAGGTGCAGTTATTTTGAACAAAGATACAAAAGTTACTTT 224
Db 1194 AGTTGAACATGAATAATGATGAGTTGAAGATTTCTACCAAGAGATATAAATGAAAAGG 1253

QY 225 CTATGATGCAAAATGAGAAAGATTTCTTCAACTCCAACTGGAGATAAAAAGTTTATTTCAGA 284
Db 1254 TCAACATGAATAATAGTAGAGTTGAAGAAATTTCTACCAGAGATGATAAAAAATGAAAAAGT 1313

QY 285 ACAAATTTAACTACAGCTAATGGAATGAGNATTTATGTAAGACAACTTTAAAAAATTT 344
Db 1314 TGAACATGAATAATAGTAGAAGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTCA 1373

QY 345 AGATGACGAGAGATATGCTATTATAGATTTAACTTATAATATGCTAAAACTGTTGAAT 404
Db 1374 ACATGAATATGATAGAGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTTGNA 1433

QY 405 TAAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTTGTTCTAGTGATGCGAAAAAATAGTGC 464
Db 1434 TGAATAGTAGAAGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTCAACATGA 1493

QY 465 AAAAGATATAGCTGAAAAATATGTTTGAAGACAAAGACTTAGAAAAATGCATAAAAAC 524
Db 1494 AATAGTAGAGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTTCAACATGAAAT 1553

QY 525 TATAATGCTCAGATTTTCAGTAAACTGATAGTTACTATCAAGTAGTTCTTTATCCAAA 584
Db 1554 AGTGAAGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTCAACATGAATAGT 1613

QY 585 AGGAAAGAGATTACAAAGTTTCTCAACTTATAGAGCTACAAATTAATGAAGGAAGTGC 644
Db 1614 AGAGGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTCAACATGAATAGT 1673

QY 645 ATATGCTAATACACAGTAAATTAATCTAATCTACTAGTAGTAAGAGTAATTTAAGAC 704
Db 1674 AGAGGTTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTCAACATGAATAGT 1733

QY 705 TGCAGTAGAAGAGTTTACAAAAATTTGAATGCTAGTTATTTCTAATCTACAACTTTAGCTGG 764
Db 1734 GGTGAAGAAATTTCTACCAGAGATATAAATGAAAAAGTCAACATGAATAGTAGAGT 1793

QY 765 TGATGACAGAAATACAAACAGCTATAGAGATAAGTAAGAAATATTAACAATTAATGATGGCGA 824
Db 1794 TGAAGAAATTTCTCCAGAAATTTGTAATGGAAGAGTACCATCACAAACAAATAACAA 1853

QY 825 GAAATCAGATCATTCAGCTGATGTTAAAGAGATGTTAAAAATGTTGTTATAGTAGTGC 884
Db 1854 T-----GAAATATTTGAAACTATAAAACAGAGAAAAAAGAAATGAAATTTAGTGTGGA 1907

QY 885 AATGCACTAGTAGAGTGGATTAGTTGCGGCTCTCTTTAGCAGCAGAGAAAAAGATGCTCCACT 944
Db 1908 AGAAAAAGCAATCCCAAGAACCCGTGGTACCTACATTTAATGAATGAATAATGAAAAAGTTAC 1967

QY 945 ATTATTAACCTCAAAAGATAAATTAGATTGTCAGTAAATCTGAATAAAGAGAGATTTT 1004
Db 1968 TCCCAAAACCATCTGAAGGTGAATCCACTAAACCATAGATTCAAAATTTAAATAGTACA 2027

QY 1005 AGACTTTAAAAACCTTCAACAGAGATTAACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAA 1064
Db 2028 AGAAAAATAAACCAATTAAGAGAAACACCGATGATGATGTCCTCAAAACATGTTGAACA 2087

QY 1065 TAGTGTATCTAAAGAAAGTTGTTAAACAGAAATTAAGAATCAATGGGATTTAAAGTTTGAAGATT 1124
Db 2088 AATATACAAGAAGATGATTAATGATGAAGAGGATGATGATGATATAGATTTTGAAGGATT 2147

QY 1125 CTCAGGTGATGATGATGATGAACTCTTTTAAATA 1161
Db 2148 ATCAAGAAAGATGATGAAAGGATTCATCAATAAAA 2184

RESULT 5

US-08-956-171E-63
; Sequence 63, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-08-956-171E-63

Query Match 4.7%; Score 85.2; DB 4; Length 8155;
Best Local Similarity 43.1%; Pred. No. 1.1e-08;
Matches 685; Conservative 0; Mismatches 883; Indels 20; Gaps 5;

QY 212 CAAAAGTTACTTTCTATGATGCAATGAGAAAGATTTCTCAACTCCACTGGAGATAAAA 271

Db 5084 CAGTAGTTACTGCAACGCTGATATAGATAAATCTGTCAGCAAAATGATGTGGATAATG 5143
Qy 272 AGTTTATTCAGAAACAACTTTAACTACAGCTTAATGGAATGAAGATTATGTAAGACAA 331
Db 5144 CAAAACTACAAATGAAGCTTACAAATCGCAGCCATTACACTGATGCAAAATGTTAAACCG 5203
Qy 332 CTTTAAAAAATTTAGATGAGGAGAAATATGCTATTATAGATTTAACTTATAATTAATGCTTA 391
Db 5204 CAGCAAAAACAGCAATTCAGATAAAGTACAGCTCAAGAAACAGCAATTTGATGGAAATA 5263
Qy 392 AAATCTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAAGAAACAGAGTAGTTGTTATCTAGTGATG 451
Db 5264 ACGGCTCAACAACTGAAGAAAAAGCAGCTGCTAAACAAACAAGTTTCAAACTGAAAAAACA 5323
Qy 452 CGAAAAATAGTGCAAAAGATATAGCTGGAAGAAATATGTTTGAAGCAAAAGCTTTAGAAA 511
Db 5324 CAGCTGATGCGCAATAGATGAGCAATCAAAATCGGAAGTTGAAGCGGGCTAAAAAAAG 5383
Qy 512 ATGCACCTAAAACTATATAATGCCTCAGATTTCAAGTAAAACTGATAGTTACTATC---AAG 568
Db 5384 CAGCAATTCGTAATTTGAAGCAATTCAGCCACAAACAACTAAAGATAATGCGAAG 5443
Qy 569 TAGTTCTTTATCCAAAGGAAAGAGATTACAAGTTTCTCAACTTATAGAGCTTACAAAT 628
Db 5444 AAGCAATTCGTAAGCAATGAAAGCGAATGAAGCTTAAACAGCAATCGCTCAAAACGCAAGACATTA 5503
Qy 629 ATAAATGAGAACTGCAATATGTTAATACAGCAGTAAATTAATCTTAAATCTTACTAGTA 688
Db 5504 CTGCTGAAGAAATTCGACGGGCTTAATCGGAGCTAGATTAATGCTGTGACACAAGCAATA 5563
Qy 689 AGAGTAATTTAAAGACTGC-----AGTAGAAGAGTTTACAAAAATGAAATGCTAGTT 739
Db 5564 GCACATTTGAAGCTGCTAATAGTCAAAATGATGTAGCAAGCGAAACGACAGGTGAAA 5623
Qy 740 ATTCTAATACTCAAACTTTAGCTGGTGATGACAGAAATPACAAACAGCTTATAGAGATAAGTA 799
Db 5624 ATAGTTATGATCAAGTAACCAACCAACAGTTAATAAAAAAGCACTGCACGCTAATGAAATCA 5683
Qy 800 AAGAAATTAACATAATATGATGGGAGAAATCAGATCATTCAGCTGATGTTTAAAGAGATG 859
Db 5684 CAGCAATTTTAAATAACAAATTTGCAAGAGATTTCAAGCTACGCCAGATGCAACAGATGAAG 5743
Qy 860 TTAAGATGTTGTTATAGTAGTAGTGCATTCACCTAGTAGATGATGATGTTGCGGCTCCTT 919
Db 5744 AAAAAACAGCAGCTGATGCTGAAGCAATTAATGAAATGTTAAAGCAATCAAGCCATTT 5803
Qy 920 TAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTCAAAAGATAAAATTAGATTCGTGAC 979
Db 5804 CAGCAGCACTACTTAACGCACAGTTGATGAAGCT---AAGCAAAATGCAGAAAGCAGCA 5860
Qy 980 TAAAACTGAAATAAAGAGAGTTTATAGACTTAAATACTCAACGAAGTAAACAGGAAAAA 1039
Db 5861 TTAATGCGGTAAACCAAAAAATTTGGAAGAAACAAAGCGGCTTAAAGATGAAATTTGATCAAT 5920
Qy 1040 CAGTTTATAGCTGGTGGAGTTAATAGTGTATCTTAAGAGTTGTAACAGAAATTTAAGAT 1099
Db 5921 TACAAGCAACGCAAAACAAATGTTATCAATATGATCAGACGCTTACAAACAGAAAGAAAG 5980
Qy 1100 CAATGGGATTTAAAAATTTGAAAGATTTCTCAGTGATGATAGATAGAACTTCTTTAAAAA 1159
Db 5981 AAGCAGCTATTCAACAATTAGCAACAGCAGTTACAGACGCGGAAAAATAATATTACAGCTG 6040
Qy 1160 TAGCAGGTGAAATAGGCTTATAGATAATGATTAAGGCTTATGTTGGTGGAGCAGGATTAG 1219
Db 6041 CAACCTGATGATAATGGTGTAGATCAGGCGAAAGACGCTGGAAGAAATTTCAATTTCAAAAGCA 6100
Qy 1220 CAGATGCCATGAGTATAGCTTCACTTCACTTAAATAGATGTTAAATGTTGTTGTTAG 1279
Db 6101 CGCAACAGCAACAGCGGTTAAATCAATGCTTAAATAATGATGTTGATCAAGCTGTGACAA 6160
Qy 1280 ATAGCAACAAATGGACATGCTACTCCAAATAGTTGTTGATGAGGAAAAAGCTGATAAAATAT 1339

Db 6161 CTCAAAATCAAGCAATTTGATTAATAACAATGGTGGTCTACAACCTGAAGAGAAAAAATGCAGCAA 6220
Qy 1340 CTGATCACTTAGATAGTTTCTTAGGAAGCGCTGATGTAGATATAATPAGGTGATTTGCAA 1399
Db 6221 AAGATTAGTTTAAAGCTAAAGAAAGCGTATCAAGATATCTTAAATGCACAAACAA 6280
Qy 1400 GTGTATCTGAAAAGATGGAAGAGCTATATCAGATGCTACTGTTAAAGCGGTTCACAGAG 1459
Db 6281 CTAATGATGTTTACGCAAAATTAAGATCAAGCAGTTGCTGATATTCAAGGTATTACTGCGAG 6340
Qy 1460 TTAAGGCGGACGATAGCAAGACACTTAACCTCTGAAGTTTATAAAAAACATATTATGCTTAATG 1519
Db 6341 ATACAAAC---AATTAAGATGTTGCAAGAGATGAATTACCAACAAAGCAACGAACAA 6397
Qy 1520 ATACTGAATAGTAAAGCTGCAGTTTATAGATAAAGATTCAAGTGTCTTCAAGTAGTGATG 1579
Db 6398 AAGCGCTTATTGCACAACTGCAGATGCGACTACTGAAGAAAAAGAAACAAGCAAAATCAAC 6457
Qy 1580 CAGGAGTATTAAATTTCTATGATGCTAAAGATGGATCTACAAAAGAGATCAATTTAGTTG 1639
Db 6458 AAGTAGACGCACAAATTAACAAGGTAAATCAAAATATTGAAAATGCAAGTCAATCGATG 6517
Qy 1640 ATGCATTAGCAGTAGGAGCTGTTGCTGGATATAAACTTCCTCAGTTGTTATAGCTTACTG 1699
Db 6518 ATGTAACACTGCAAAAGATAATGCAATTCAGCAATTCACCCANTTCAGCATCAACAG 6577
Qy 1700 ATTCTTTATCTTCTGATCAATCGTTGCTTATAGCAAAAGTTGTAGAGAAAAATATTCTTA 1759
Db 6578 ATGTTAAACGATGCAAGCGGAATTCGTTAACTGAATGCAA--AATAAAATAACTGA 6635
Qy 1760 AAGATTTAACACAAGTTGTTGTCAGGAAT 1787
Db 6636 AATACTTAATAATAGAGACTACTAAT 6663

RESULT 6

US-08-781-986A-63
; Sequence 63, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: *Staphylococcus aureus* Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8155 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-63

Query Match 4.7%; Score 85.2; DB 4; Length 8155;
Best Local Similarity 43.1%; Pred. No. 1.1e-08;
Matches 685; Conservative 0; Mismatches 883; Indels 20; Gaps 5;

QY 212 CAAAGTTACTTCTATGATGCAAAATGAGAAAGATCTTCAACTCCAACTGGAGATAAAA 271
DB CAGTAGTTACTGCAACCGCTGATATAGATAATGCTGCAGCAAAACAATGATGGGATAATG 5143

QY 272 AAGTTTATTCAGAACAACTTTAACTACAGCTTAATGGAATCAAGATTTGTAAGACAA 331
DB CAAAACTACAAATGAAGTCACTCAATCGCAGCCATTACACTGATGCAAAATGTTAAACCGAG 5203

QY 332 CTTTAAAAAATTTAGATGCGAGAGAAATATGCTATTATAGATTAACTTTATATAATATGCTTA 391
DB CAGCAAAAACAGCAATTTGAGATAAAGTACAAGCTCAAGAAACGCAATTTGATGGAAATA 5263

QY 392 AAACCTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAAGAAAACAGTAGTTGTTATCTAGTGATG 451
DB ACGGCTCAACAACTGAAGAAAAGCAGCTGCTTAAACAACTGTTCAAACTGAAAAAACA 5323

QY 452 CGAAAAATAGTCAAAAGATATAGCTGAAAAATATGTTTGAAGCAAGAACTTAGAAA 511
DB CAGCTGATCGCCCAATAGATGAGCAACATACAAATGCGGAAGTTGAAGCGGCTTAAAAAG 5383

QY 512 ATGCACTAAAAACTATAAATGCTCAGATTTTCAGTAAAACTGATAGTTACTATC---AAG 568
DB CAGCAATTTGCTAAATTTGAAGCGATTGAGCAAGCACTCAACTAAAGTAATGCGAAAG 5443

QY 569 TAGTCTTTTATCCAAAGGAAAGAGATTAACAAGTTTCTCACTTATAGAGCTACAAATTT 628
DB AAGCAATTTGCTAGAAAGCGAATGAACGTTAAACAGCAATCGCTCAACGCAAGACATTA 5503

QY 629 ATAAAGGAAGTGCATATGTTAATACACCGATTAATTAATCTAAATCTACTAGTA 688
DB CTGCTGAAGAAATTTGAGCGGCTTAATGCGGACGTAGATAATGCTGTGACACAAAGCAATA 5563

QY 689 AGAGTAATTTAAAGACTGC-----AGTAGAGAGTTTACAAAATTTGAATGCTAGTT 739
DB GCACATTTGAGCTGCTAATAGTCAAAATGATGAGCAAGCAAGCAAGTGAA 5623

QY 740 ATTCTAATCTACAACTTTAGTGGTGATGACAGATAACAAACAGCTATAGAGATAAGTA 799
DB ATAGTATTGATCAAGTAAACCAACAGTTAATATAAAAAAGCACTGCACGTAATGAAATCA 5683

QY 800 AAGAAATTACAAATATGATGGGAGAAATCAGATCATTTCACTGATGTTTAAAGAGAAATG 859
DB CAGCAATTTTAAATAACAAATTTGCAAGAGATTTCAAGCTAGCCAGATGCAACAGATGAAG 5743

QY 860 TTAATAATGTTGATTAGTAGGTGCAAAATGCACTAGTAGATGATTTAGTTGCGGCTCCCTT 919
DB AAAAAACAGCAGCTGATGTAAGCAAAATCTGAAATGTTAAAGCAATCAAGCAATTT 5803

QY 920 TAGCAGCAGAAAAAGATGCTCCACTATTATTAATCTTCAAAAGATAAATTTAGATTCGTACG 979
DB CAGCAACAACTACTACGCAAGATTTGATGNAAGCT---AAAGCAATGCAAGACGACGCA 5860

QY 980 TAAATCTGAAATTAAGAGAGTTTATAGCTTAAAACTTCAACAGAGAGTAAACAGGAAAA 1039
DB TTAATGCGGTAAACCAAAAGTTGTGAAGAAACAAAGCGCTAAAGATGAAATTTGATCAAT 5920

QY 1040 CAGTTTATATAGCTGGTGGATTAATAGTGTATCTAAAGAGTTTGAACAGAAATTAGAAT 1099
DB TACAAGCAACGCAAAACAATGTTTATCAATAATGATCAGAACGCTACACAGAGAAAAAG 5980

QY 1100 CAAATGGGATTTAAAGTTGAAAGATTTCTCAGGTGATGATAGATATGAATCTTTTAAAAA 1159
DB AAGCAGCTATTCAACAAATTAGCAACAGCAGTTTACAGACGCGAAAAAATAATTTTACAGCTG 6040

QY 1160 TAGCAGGTGAAATAGGCTTAGATAATGATAAGGCTTTATGTTGTTGGTGAACAGGATTAG 1219

DB 6041 CAACGTGATTAATGGTGTAGATCAGGCGAAGACGCTGGAAAGAAATTCATTTCAAAGCA 6100

QY 1220 CAGATGCCATGAGTATAGCTTTCAGTTGCTTCTCTAAATTAGATGGTAATGGTGTGTAG 1279

DB 6101 CGCAACCAAGCAACAGCGGTTAAATCAAAATGCTAAAAATGATGTTGATCAAGCTGTGACAA 6160

QY 1280 ATAGAACAATGGACATGCTACTCCAATAGTTGTTGTAGATGGAAAGAGCTGATAAATAT 1339

DB 6161 CTCAAAATCAAGCAATTTGATAATCAACTGGTGTCTCAACTGAAGAGAAAAAATGCAGCAA 6220

QY 1340 CTGATGACTTAGATAGTATTTCTTAGAAGCGCTGATAGATATATAATAGTGGATTTGCAA 1399

DB 6221 AAGATTAGTTTTTAAAGCTTAAAGAAAGGGTATCAAGATATCTTAAATGCAACAACAA 6280

QY 1400 GTGTATCTGAAAAGATGGAAGAGCTATATCAGATGCTACTGGTAAAGGCGTTTCAAGAG 1459

DB 6281 CTAAATGATGTTTACGCAAAATTAAGATCAAGCAGTTGCTGATATTTCAAGGTATTACTGCAG 6340

QY 1460 TTAAGGCGCGATAGACAGACACTAACTCTGAAGTTTATAAACAATATTTATGCTAATG 1519

DB 6341 ATACAAC---AATTAAGATGTTGCGAAGATGAAATTAGCAACAAAAGCAACGAAACAA 6397

QY 1520 ATACTGAAATAGCTAAAGCTGCAGTTTATAGATAAAGATTCAAGTGTCTTCAAGTAGTGATG 1579

DB 6398 AAGCGCTTTTTCACAACTGCAATGCGACTACTGAAGAAAGAAAGCAACAATCAAC 6457

QY 1580 CAGGAGTATTTAAATTTCTATGATGCTAAGATGAGTCTACAAAAGAGATCAATTTAGTTG 1639

DB 6458 AAGTAGCGCAATTTAAACACAGGTAAATCAAAATATTTGAAAAATGCAACAGTCAATCGATG 6517

QY 1640 ATGCATTAGCAGTAGGAGCTGTTGCTGATATAAATCTGCTCCAGTTGTTATTTAGTACTG 1699

DB 6518 ATGTAACACACTGCAAAAGATAATGCAATTCAGCAATTTGACCCCAATTCAGCATCAACAG 6577

QY 1700 ATCTTTTATCTTCTGATCAATCGGTTGCTATAAGCAAAAGTTGTAGGAGAAAAATATTCTA 1759

DB 6578 ATGTTAAACCAATGCAAGACGCGAATTTGCTTAACCTGAATGCAA--ATAAATAACTGA 6635

QY 1760 AAGATTTAACACAAAGTTGCTCAAGGAAT 1787

DB 6636 AATACTTAATAATGAGACTACTAAT 6663

RESULT 7
US-09-754-947-2
; Sequence 2, Application US/09754947
; Patent No. 6828110
; GENERAL INFORMATION:
; APPLICANT: Lee, Bruce Andrew
; APPLICANT: Flores, Becky Mar
; APPLICANT: Valiure, Gunars Edwin
; APPLICANT: Biosite Diagnostics, Inc.
; TITLE OF INVENTION: Assays for Detection of Bacillus Anthracis
; FILE REFERENCE: 014907-003310US
; CURRENT APPLICATION NUMBER: US/09/754,947
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,901
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; FEATURE:
; OTHER INFORMATION: surface array protein (SAP)
US-09-754-947-2

Query Match 4.5%; Score 82.2; DB 4; Length 2370;
Best Local Similarity 42.3%; Pred. No. 3.5e-08;
Matches 715; Conservative 0; Mismatches 958; Indels 18; Gaps 4;

Qy	2	TGAAAAAGAAATTTAGCAATGGCTATGGCAGCTGTTACTGTAGTAGTGGTTCTGCTGCTC	61
Db	668	TTAAAGAGGTAACTTTTATCAGAAGATAAAAGATCTGCTACAGTTGGAATATATATAGTAACT	727
Qy	62	CAGTTTTTGCAGCAGAGCTTCAGATGTAATATACACTACAAGATGTTACAAATGATAAGTATA	121
Db	728	TAGAGCTTAAACAAACTTACACTGTAGATGTAACAAAGTTTGGTTAAACACAGAAGTAGCTG	787
Qy	122	CAGTATCCAAATACCTAAGAGCTAGTGACTTAGTAAGAGGATATTTTAGCAGCACAAACTTAA	181
Db	788	TAGGTTCTTTAGGAAGCAAAAACAATCGAATGGCTGACCCAAACAGTTGTAGCTGATGAGC	847
Qy	182	CAACAGGTGCAGTTATTTTGGAAACAAGATACAAAAGTTACTTCTTATGATGCAAAATGAGA	241
Db	848	CAACAGCATTTACAAATTCACAGTTTAAAGATGAAAACGGTACTGAAGTTGTTTCCACGAGG	907
Qy	242	AAGATTCTTCAACTCCAACTGGAGATAAAAAGTTTATTTCAGAACCAAACTTTTAACTACAG	301
Db	908	GTATTTGAATTTGTAAACGCCAGCTGCGAGAAAAAATTAATGCAAAAGGTGAATACCTTTAG	967
Qy	302	CTAATGGAAATGAAGATTATGTATAAGACAACCTTTTAAAAAATTTAGATGCAAGGAAATATG	361
Db	968	CAAAAGGTACTTCAACTACTGTATAAAGCTGTTTATAAAAAGACGGTAAAGTAGTAGCTG	1027
Qy	362	CTATTTAAGATTTAACTTTATAAATAGCTTAAAACTGTTTGAATTTAAAGTAGTAGCAGCTA	421
Db	1028	AAAGTAAAGAAGTA-----AAAGTTTCTGCTGAAGGTGCTGCGCAGTTCCTCAATCT	1078
Qy	422	GTGAAAAAAACAGTAGTTGTATCTAGTGCATCGGAAAAATAGTGCAAAAGATATAGCTGAAA	481
Db	1079	CTAACTGGACAGTTTGCAGAACAAAATAAAGCTGACTTTTACTTCTTAAAGATTTCAAACAAA	1138
Qy	482	AAATATGTGTTTGAAGACAAAAGCTTAGAAAATGCATTAATAAACTATAAAATGCGCTCAGATT	541
Db	1139	ACATAAAGTTTACGAAGCGCAACAACGCTTACGTTCAAGTAGAATTTGAAGATCAATTTTA	1198
Qy	542	TCAGTAAAACTGATPAGTTACTATCAAGTAGTCTTTTATCCAAAAGAAAGAGATTTACAAG	601
Db	1199	ACGCAGTAAACAACCTGGAAAAAGTTGCAATATGAGTCGTTTAAAAACACAGAAGTTGCTGTAGTAG	1258
Qy	602	GTTTCTCAACTTATAGAGCTACAAATTTATAATGAGAGAACTGCATATGGTAATACACCCAG	661
Db	1259	ATAAAGCTACTGGTAAAGTAACTGTATATCTGCAGGAAAAGCCACAGTAAAGTAACTG	1318
Qy	662	TAATA-----TTAACTCTAAAACTACTAGTAAAGAGTAATTTAAAGACTGCGAGTAGAAGAGT	718
Db	1319	TAAAGAATTCAAAAGGTAAAGCACTGTTTTCACACACAGTTGAAATTTGAAGCTTTGCGCTC	1378
Qy	719	TACAAAAATGAAATGCTAGTTATTTCTTAATACTACAACTTTTACGCTGGTGATGACAGAAATC	778
Db	1379	AAAAAGCAATGAAAAGACATTAATTTAGAAAAAACTAACGTAGCGCTTTCTACAAAAGATG	1438
Qy	779	AAACAGCTATAGATAAGTAAAGAAATTTACAAATATGATGGCGAGAAATCAGATCAATT	838
Db	1439	TAAACAGATTTAAAAAGTAAAAAGCTCCAGTACTAGATCAATAACGTTAAAGAGTTTACAGCTC	1498
Qy	839	CAGCTGATGTTTAAAGAGAACTCTTAAAAATGTTGTATTTAGTAGGTGCGAAATGCACACTAG	898
Db	1499	CTGTACAGTGAAGTACTTGATAAGAGTGTAAAGAAATTAAGAACAATAATTAGAG	1558
Qy	899	ATGATTAGTTGCGCGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTCAA	958
Db	1559	CTAAATATGTGAACAGAGAAATTAGTCTGTAATGCGAGCGTCAAGAACGTTGTAATTATA	1618
Qy	959	AAGATAAATTTAGATTGTCAGTAAAAATCTGAAATTAAGAGAGTTTTTATAGACTTTAAAAACTT	1018
Db	1619	CAGTTGTATTAACCTGCAAAATCTCGTGAAAAAGAAAGCAAAAGCTACATTAGCTCTAGAAT	1678
Qy	1019	CAACAGAAGTAAACAGGAAAAACAGTTTATATAGCTGTTGGAGTTTAAATAGTGTATCTAAAG	1078
Db	1679	TAAAGCTCCAGGTGCATTTCTCTAAATTTTGAAGTTTCGTGGTTTAGACACAGAAATTAGATA	1738
Qy	1079	AAGTTGTGAACAGAAATTAGAATCAATGGGATTTAAAAGTTTGAAGAAATTTCTCAGGTGATGATA	1138

Db	1739	AAATATGTTCTGAGGAAACCAAGAAATGCAATGACTGTTTCAGTTCCTCTGTAGATG	1798
Qy	1139	GATATGAAACTTCTTTAAAAATAGCAGGTGAAATAGCTTAGTAATAGTAAGGCTTATG	1198
Db	1799	CAAAATGGATTAGTATTAAAGGTGCGAAGAGCAGCTGAACCTAAAAGTAACAACAACAACA	1858
Qy	1199	TAGTTGGTGAACAGAGATTAGCAGATGCCATGAGTATAGCTTCAGTTGCTTCTACTAAAT	1258
Db	1859	AAGAAGT---AAAGAAGTAGACCAACTGATGCCAAGTTACTGTACAAAAATAACAGTG	1915
Qy	1259	TAGATGGTAATGCTGTTGTAGATAGAAACAAATGGACATGCTACTCCAAATAGTTGTTGTAG	1318
Db	1916	TAATTACTGTGTGCTCAAGGTGCAAAAGCTGGTGAGACTTATAAAGTAACACAGTTGTACTAG	1975
Qy	1319	ATCGAAAAGCTGATAAAATATCTGATGACCTTAGATAGTTCTTTPAGGAAGCGCTGATGTAG	1378
Db	1976	ATGGTAAATTAATCAACAATCATTCATTCAAAGTTGTTGATACAGCACCAACTGCTAAAG	2035
Qy	1379	ATATAATAGGTGGAATTTGCAAGTGTAATCTGAAAAGAGATGGAAGAGCTPATATCAGATGCTTA	1438
Db	2036	GATTAGCAGTAGAATTTACAAGCACACTCTTAAAGAAAGTAGCTCCAAATGCTGATTTAA	2095
Qy	1439	CTGGTAAGCGTTACAGAGTTTAAGGGCAGCATAGACAAGACACTAACTCTGAGTTA	1498
Db	2096	AAGCTGCACCTTTTAAATATCTTATCTGTGGTGGTACCTGCGACTACAGCAAAAGCAA	2155
Qy	1499	TAAAAACATATTATGCTTAATGATACTGMAATAGCTAAAGCTGCAGTCTTTTAGATAAAGATT	1558
Db	2156	CAGCTTCTAATGTAGAATTTGTTTCTGCTGCACACAAATGTTGTAGCT---GAAAAATGGTA	2212
Qy	1559	CAGGTGCTTCAAGTAGTGATGCGAGGAGTATTTTAATTTTCTATGTAGCTAAAGATGGATCTA	1618
Db	2213	CAGTTGGTGCAAAAGGTGCAACATCTATCTATGTGAAAAACCTGCACAGTTGTTAAAAGATG	2272
Qy	1619	CAAAAAGAGATCAATTAGTTGTATGTCATTAGCAGTAGGAGCTGTTGCTGGGATATAAAACTTG	1678
Db	2273	GAAGAAGCAAAAGTAGAATTTGATAAAGCTGTACAAAGTTGCAGTTCTTATTAAAGAAG	2332
Qy	1679	CTCCAGTTGTA	1689
Db	2333	CAAAACCTGCA	2343

RESULT 8

```

US-09-902-540-1280
; Sequence 1280, Application US/0902540
; Patent No. 683347
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

```

Query Match 4.2%; Score 76; DB 4; Length 1039;
Best Local Similarity 44.2%; Pred. No. 5.8e-07;
Matches 408; Conservative 0; Mismatches 507; Indels

2:

```
QY 267 TAAAAAGTTTATTTCAGACCAACTTTAACTACAGCTAATGGAATGAGATTATGTAAA 326
Db 69 TAAAAACCAACCAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 128
QY 327 GACAACTTTAAAAATTTAGATGCGAGAGAAATATGCTATTATAGATTAACTTTATAATAA 386
Db 129 AAAATAAAAAACAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 188
QY 387 TCGTAAAACTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAAGAAAAACAGTAGTTGTACTAG 446
Db 189 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 248
QY 447 TGATGCGAAATAGTGCAGAGATATAGCTGAGAAATATGCTTTGAGACAAAGACTT 506
Db 249 AAAAAACAATACATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 308
QY 507 AGAAATGCACTAAAAACTATAAATGCCCTCAGATTTCAGTAAAACTGATAGTTACTATCA 566
Db 309 ATAAAAAANA -AAAAATTAATATCAATCATATAAATAAAAAATAAAAAATAAACA 366
QY 567 AGTAGTTCTTTATCCAAAGGAAGAGATTAACAAGTTTCTCAACTTATAGAGCTACAAA 626
Db 367 AAAAAACCAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 426
QY 627 TTATATGAGGACTGTCATATGTTAATACACAGCTAATATTAACTCTAAATCTACTAG 686
Db 427 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 486
QY 687 TAAAGTAGTAAATTAAGACTGCGAGTAGAGAGTTTACAAAAATTTGAATGCTAGTTATTCTAA 746
Db 487 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 546
QY 747 TACTACAACCTTTAGCTGGTGATGACAGAAATACAAACAGCTATAGAGATAAGTAAGAATA 806
Db 547 TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 606
QY 807 TTACATATATGATGCGAGAAATCAGATCATTCAGCTGATGTTAAGAGAAATGTTAAAAA 866
Db 607 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 666
QY 867 TGTGTATTAGTAGTGCAGATGCTAGTAGATGATTTAGTTGGGGCTCCTTTAGCAGC 926
Db 667 TGTTTTTT-----CTTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 719
QY 927 AGAAAGATGCTCCATTTATTAACCTTCAAAAGATAAATTTAGATTTCTCAGTAAATC 986
Db 720 AAGAAATTAACAAAGGCTCTAATCAACACCCCAATATACAAAAATAATCTTAAAAACAAA 779
QY 987 TGAATTAAGAGAGTTTATAGACTTAAACCTTCAACAGAGTAAACAGGAAACAGTTTA 1046
Db 780 AAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 839
QY 1047 TATAGCTGGTGGATTAATAGTGTATCTAAAGAGTTGTAACAGAAATTAGAATCAATGG 1106
Db 840 AGGGAATAAAAAACAATAAATACTTAAAAATAAAAAATAAAAAATAAAAAATA 899
QY 1107 ATTAAAAAGTTGAAGATTTCTCAGGTGATGATAGATGAACTCTTTAAAAAATAGCAGG 1166
Db 900 AATAAAGAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 959
QY 1167 TGAATAGGCTTAGATTAATGATAA 1190
Db 960 AAAAAAATAAAAAATAAAAAATAA 983
```

RESULT 9

PCT-US93-07261-10

; Sequence 10, Application PC/TUS9307261

; GENERAL INFORMATION:

; TITLE OF INVENTION: PfEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

```
; ADDRESSEE: John H. C. Blasdale
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940-1000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07261
; FILING DATE: 19930805
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/927,531
; FILING DATE: 07-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blasdale, John H. C.
; REGISTRATION NUMBER: 31,895
; REFERENCE/DOCKET NUMBER: DX0288K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-822-7398
; TELEFAX: 201-822-7039
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4766 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; STRAIN: Malayan Camp
; IMMEDIATE SOURCE:
; CLONE: p2b1;p12-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..4766
; PCT-US93-07261-10
```

Query Match 4.0%; Score 73.4; DB 5; Length 4766;

Best Local Similarity 42.4%; Pred. No. 3.2e-06;

Matches 651; Conservative 0; Mismatches 871; Indels 15; Gaps 4;

QY 37 GTTACTGTAGTAGTTCTGCTGCCAGTTTTTGCGAGAGCTTCAGATGTATATACATA 96

Db 2124 GGTAAAGGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2183

QY 97 CAAGATGCTACAAATGATAGTATACAGTATCAATATCTAAAGCTAGTGTACTAGTAAG 156

Db 2184 GCAAAAGGTAAGGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2243

QY 157 GATATTTTTCAGAGCAACAACTTAAACACAGCTGCAGTTATTTTGAACAAAGATACAAA 216

Db 2244 GATTTCAGAAAGGTAAGGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2303

QY 217 GTTACTTCTATGATGCAAAATGAGAAAGATTTCTTCAACTCCAACTGGAGAGATAAAAAAGTT 276

Db 2304 GAAACCGATTTCAGCAAAAGGTAAGGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2363

QY 277 TATTCAGAACAACTTTTAACTACAGCTAATGGAATGAAGATTATGTAAGACACACTTTA 336

Db 2364 TATATGAACCTGATTTTAGCAAAAGGTAAGGAGTAAGTAAGTAAGTAAGTAAGTAAG 2423

QY 337 AAAAATTTTAGATGCGAGAGAAATATGCTA---TTATAGATTTTAACTTATAATAATGCTAAA 393

Db 2424 GAAGATATATGAACTGATTTAGCAAAAGGTAAGGAGTAAGTAAGTAAGTAAGTAAGTAAG 2483

QY 394 ACTGTTGAAATTAAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 453

Db 2484 AATCTAGGAAGTAATATGAACTGATTTAGCAAAAGGTAAGGAGTAAGTAAGTAAGTAAGTAAG 2543

454 ABAATAGTGCRAAGATATAGCTGAAATAATGTGTTGAAGACAAAGACTTAA---GAA 510
 2544 CGCGAAATCTAGAGAAATATATGAACTGATTTAGCAAAAGGTAAAGGATTAATAT 2603
 511 AATGCCTACTAAAACTATAATGCTCAGATTTTCAGTAAAACTGATAGTTACTATCAAGTA 570
 2604 AAGACACGCGAAATCTAGAGAAATATATGAACTGATTTAGCAAAAGGTAAAGGATTA 2663
 571 GTTCTTTATCCAAAAGAAAGAGATTACAAGGTTTCTCAACTATATAGAGCTCAAAATTA 630
 2664 ACTAATAAAGACGCGAAATCTAGAGAAATATATGAACTGATTTAGCAAAAGGTAAAG 2723
 631 AATGAAGGAACTGCATATGTAATACACAGTAAATTAATCACTCAAAATCTACTAGTAAG 690
 2724 GAAGTAACTAATAAGCAGCGAAATCTAGAGAAATATATGAACTGATTTAGCAAAAGGT 2783
 691 AGTAAATTTAAGAGCTGCAGTAGAAGATTACAAAAATTTGAATGCTAGTTATTTCTAAT 750
 2784 GGTAAGGAGTAACTAATAAGCAGCGAAATCTAGAGAAATATGAACTGATTTAGCAAAAG 2843
 751 ACACTTTAGCTGCTGATGACAGAAATACAAACAGCTATAGAGATTAAGGATTAATAC 810
 2844 ATGAAAAATTAATGAATTTACAAATTAAGGATCTGATGATTAAGGAAATGAGAGCTA 2903
 811 AATAATGATGCGAGAAATCAGATCAATTCAGCTGATGTTAAAGAGAAATGTTAAAAAT 870
 2904 AAAAAATAAGAAATTAAGAAATTAAGGATCTGATGATTAAGGAAATGAGAGCTAATA 2963
 871 GTATTAGTGTGCAATGCACTAGTAGATGATGATTTAGTTCGGCTCTTTAGCAGAGAA 930
 2964 AATAAGAAATTAAGAAATTAAGGATCTGATGATTAAGGAAATGAGAGCTAATAATA 3023
 931 AAGATGCTCCTATTAATTAATCTCAAAAGATTAATTAATGATTCGTCAGTAAATCTGAA 990
 3024 AAGAAATTAAGAAATTAAGGATCTGAGGATTAAGGAAATTAAGGAAATTAAGAAATA 3083
 991 ATAAGAGAGTTTATAGCTTTAAACCTTCAACAGAGTAAAGGAAATTAAGGAAATTA 1050
 3084 GAATTACGAATTAAGGATCTGAGGATTAAGGAAATTAAGGAAATTAAGGAAATTAAG 3143
 1051 GCTGCTGAGTTAATAGTGTATCTAAAGAAATTTGTAACAGAAATTAAGGAAATTAAG 1110
 3144 TTACGAATTAAGGATCTGAGGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGAA 3203
 1111 AAGTTGAAAGATTTCTCAGGTGATGATGATTAAGGAAATTTAAAGAAATTAAGGAAAT 1170
 3204 CAAATTAAGGATCTGAGGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTA 3263
 1171 ATAGCTTAGATATGATAGGCTTTATGTTGTTGTAAGGAAATTAAGGAAATTAAGGAA 1230
 3264 AATAAGGATCTGAGGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAG 3323
 1231 AGTATAGCTTCAAGTTGCTTCTACTAAATTTAGATGATGATGTTGTTGATGATGATTA 1290
 3324 AAGGATCTGAGGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAA 3383
 1291 GGCATGCTACTCCAAATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1350
 3384 GGCATGCTGATGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAA 3443
 1351 ---GATAGTTTCTTAGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1407
 3444 TCTGATGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAG 3503
 1408 GAAAGATGGAAGAGCTTATATCAGATGCTACTGCT-----AAGGCGTTTACAAGATT 1461
 3504 GAAGATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGGAA 3563
 1462 AAGGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1521
 3564 AAGATTTATCTAATAAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 3623
 1522 ACTGAAATAGCTTAAGCTGCGATTTTAGATTAAGATT 1558

Db 3624 AAGACATGACGAATAAAGAACTATTAAATAAAGACT 3660

RESULT 10

US-09-601-198-56/c
 ; Sequence 56, Application US/09601198
 ; Patent No. 6531583
 ; GENERAL INFORMATION:
 ; APPLICANT: Cassell, Gail H.
 ; APPLICANT: Chen, Ellison Y.
 ; APPLICANT: Glass, Jennifer S.
 ; APPLICANT: Glass, John I.
 ; APPLICANT: Heiner, Cheryl R.
 ; APPLICANT: Lefkowitz, Elliot
 ; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
 ; TITLE OF INVENTION: UREALYTICUM
 ; FILE REFERENCE: UAB-13452/22
 ; CURRENT APPLICATION NUMBER: US/09/601,198
 ; CURRENT FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: 60/073,189
 ; PRIOR FILING DATE: 1998-01-30
 ; NUMBER OF SEQ ID NOS: 181
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 56
 ; LENGTH: 14066
 ; TYPE: DNA
 ; ORGANISM: Ureaplasma urealyticum
 ; US-09-601-198-56

Query Match 3.9%; Score 71.6; DB 4; Length 14066;
 Best Local Similarity 43.7%; Pred. No. 1e-05;
 Matches 537; Conservative 0; Mismatches 669; Indels 24; Gaps 4;

QY 124 GTATCAAACTAATAAGCTAGTACCTAGTAAAGGATTTTACAGACACAAAACTTAA 183
 Db 12584 GCATTAATAATAAACAGTAATTAATCCTAGACAAATTTGTTTCAGTTCAAACCCCAA 12525
 QY 184 ACAGGTGCGAGTT---ATTTTGAACAAAGATACAAAAAGTTACTTTCTATGATGCAATGAG 240
 Db 12524 AATGTTTGGTTTCAATTTTGATTTTAACTTCAAAACAAAAACAATTAATTAAGAGCCA 12465
 QY 241 AAGAGTTCTTCAACTCAACTGAGATATAAAGTTTATTCAGAACCAACTTTAACTTACA 300
 Db 12464 GCTGCTAATGCAAGTGTGATTCACCTTAAACCAACCCCAATTTTAAAGAAATTAAT 12405
 QY 301 GCTAATGGAATCAAGATTAATGTAAGACAACTTTTAAAGAAATTTAGATGAGGAGATAT 360
 Db 12404 GATTTAATTAATCAGAAATTTAGTAGCAACTTTTGTGATTAATTAATGAAGAATATA 12345
 QY 361 GCTATTATAGATTTAACTTATAATAATGCTTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT 420
 Db 12344 GTAGTTGCTAATGTTGGATCAAAATTAATAAATTTATTTTGTAGTAAATGATTTTACCA --- 12288
 QY 421 AGTGAAAAACAGTAGTTGTTATCTAGTATGCGAAAAATAGTGCAAAAGATATAGCTGAA 480
 Db 12287 AAGGATATATTATCATTTTAGCTTAAAGTTGAAAAATACGATTTTAAATTAAGTCATTAAT 12228
 QY 481 AATATGTTGTTGAGCAAGCAAGCTTAGAAAAATGCACTTAAACCTATAAATGCTCAGAT 540
 Db 12227 CTTAATGATTTTGAACCTTAAAGATTAATAATTTGATTAACAGATTTGAATTTATTAAT 12168
 QY 541 TTCAGTAAACTGATAGTACTATCAAGTAGTTCTTTTATCCAAAGGAAAGAGATTACAA 600
 Db 12167 TCATCCAGATTTTAGTACGATTAATGATGGAATTTAGAAATTCACACTCAATTAGCA 12108
 QY 601 GGTTCCTCAACTTAT-----AGAGCTACAAATTAATTAAGGAACTGCATATGTTAAT 654
 Db 12107 AATGATTTAAATGATGATTTTAAACAAAAAGCTTTTAAATTAATGCAAAATGTCAGAGGAA 12048
 QY 655 ACACGATTAATTAATCACTTAATCTACTAGTAGAGATTAATTTAAAGACTGCAAGTACAA 714
 Db 12047 GTTCTTGTCAAGATGTTGATATAGAACATGAGATAGATGTAAGTATTTGATGCTTAATGGCAAG 11988

```

QY 715 GAGTTACAAAATTGATGCTAGTATTCTTAATACATCAAACTTTAGCTGGTGATGACAGA 774
Db 11987 GTTATTATTCCACAAAATTTAGCAATAATGACCCAACTAAACCTTAATATATACACA 11928
QY 775 ATACAAAACAGCTATAGAGATAAGTAAAGAAATATTACAATAATGATGGCGAGAAATCAGAT 834
Db 11927 CTTAAAAAGTCGTTTTAAAAACAAATATCAACCAATATTGATCTAATTAAGTGAAGAG 11868
QY 835 CATTGAGCTGATGTTTAAAGAGAAATGTTAAAAATGTTGTTATTAGTAGGTCNAATGCACATA 894
Db 11867 CAATTAAAGTGGTGATAATCATATTTTCATTTAAAAAACCAACAATAACTGCTAAAACTAAA 11808
QY 895 GTAGATGGAATAGTTCGGCTCTTTAGCAGCAGAGAAAAGATGCTCCACTATTTAACT 954
Db 11807 GAAAATGATGATTAATGAAATTTAGTTTTTCA-----AATCCTAGTTTTAGCAAT 11760
QY 955 TCAAAAGATAAATTAGATTCGTCAGTAAATCTGAAATAAAGAGAGTTTTAGACTTTAAAA 1014
Db 11759 AAAAAAATTAAATTAACATTTTAAACCGATGATTAATACTAATAACAAAAACGGTTGAA 11700
QY 1015 ACTTCAACAGAGATTAACAGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGTATCT 1074
Db 11699 GCTAGTATTGGTTAGATGGAAGAGCTATTTTAAAAACAAGTGATGATCAATTTTGTCA 11640
QY 1075 AAGAAGTTGTAACAGAAATAGAAATCAATGGATTAAGATTTAAAGTTGAAGATTCACAGTAT 1134
Db 11639 CCAGATCATAAATATATACACTACAAAATTTGAAGCTGATTAATAAAAAAGTTGCTTAAT 11580
QY 1135 GATAGATATGAACCTCTTTAAAAATAGCAGGTGAATAGGCTTAGATAATGATGAAGCT 1194
Db 11579 GATGAATTTCCATTAGATGATGATGATTAATAACAAAAATGTTAATGTTAATGCT 11520
QY 1195 TATGTAGTTGGTGGAAACAGATTTAGCAGATGCGATGATGATGATGATGATGATGATGAT 1254
Db 11519 GATAATAAGCATGAGTTTAAATATACCAGATCAAAAAATAAAGATTTAACTGCTGTTAT 11460
QY 1255 AAATTAGATGGAATGTTGTTAGATAGAAATGATGATGATGATGATGATGATGATGATGAT 1314
Db 11459 AAGATATAAATAAACAATGAAATTCATGTTCCAAATCAAACTGATGATGAAGGAAAGTT 11400
QY 1315 GTAGATGAAAAGCTGATATAATATCTGAT 1344
Db 11399 ATGTTAATCCAAACATTAATTTTGAT 11370

```

```

RESULT 11
US-08-446-137B-1
; Sequence 1, Application US/08446137B
; Patent No. 6162903
; GENERAL INFORMATION:
; APPLICANT: Trowern, Angus R.
; APPLICANT: Atkinson, Anthony
; APPLICANT: Murphy, Jonathan P.
; APPLICANT: Laurence, Oliver S.
; APPLICANT: Dugglesby, Clive J.
; TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
; FROM L PROTEIN AND THEIR USES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,137B

```

```

Query Match 3.8%; Score 70; DB 3; Length 3279;
Best Local Similarity 42.8%; Pred. No. 1.5e-05;
Matches 468; Conservative 0; Mismatches 620; Indels 6; Gaps 2;

QY 99 AGATGGTACAAATGATAAGTATACAGTATCAATTAAGCTAGTACTAGTAAAGGA 158
Db 1374 AAAAGAAACACCAAGAAACCCAGAGAAACCAAAAGAAAGTACAAATCAAAAGTTAACTT 1433
QY 159 TATTTTAGCAGCACAAAACCTTAAACACAGGTGTCAGTTATTTTGAACAAAGATACAAAAGT 218
Db 1434 AATCTTTCGATGGAAGATACAAACAGCAGAAATTTCAAGGAACATTTGAAGAGCAAC 1493
QY 219 TACTTTTCTATGATGCAAAATGAGAAAGATTTCTTCAACTCCAACTGGAGATAAAAAAGTTTA 278
Db 1494 AGCAAAAGCTTATGCTTATGCAAACTTATAGCAAAAGAAATGCGAATATACAGCAGA 1553
QY 279 TTCAGAACAAACTTTAACTAGCTAGCTAATGGAATGAAGATTAATGAAGACAACTTTAA 338
Db 1554 CTTAAGATGGTGGAAACACAACTTAATTTGCTGGAAGAAACACCAAGAAC 1613
QY 339 AAATTTAGATGACGAGAAATATGCTTATTAATAGATTTAACTTATAATAATGCTAAAACTGT 398
Db 1614 ACCAGAAACCAAAAGAAAGAGTTACAACTCAAGTTAACTTAACTTTGCAGA---TGG 1670
QY 399 TGAATTAAGTAGTAGCAGCTAGTGAAAAACAGTAGTTGTATCTTAGTATGCGAAAAA 458
Db 1671 AAAACACAAACAGCAGAAATTTCAAGGAACATTTTGAAGAAACAAACAGCAGAAAGCTTACAG 1730

```

```

US-08-446-137B-1

```


QY 459 TAGTGCAAAAGATATAGCTCAAAAAATATGTGTTTGAAGACAAAGACTTAGAAAAATGCACCT 518
Db 1731 ATATGCGACCTTATTAGCAAAAGTAAATGTGTAATACACAGCAGACTTAGAGATGGCGG 1790
QY 519 AAAAACTATATAATGCTCAGATTTCAGTAAACCTGATAGTTACTATCAAGTAGTCTTTTA 578
Db 1791 ATACACTATCAACATCAAAATTTGCTGMAAAGAACCAACAGCGGAAAAATCCAGGAATCAC 1850
QY 579 TCCAAAAGGAAGAGATTACAAAGTTTCTCACTTATAGAGCTACAAATATATATCAAGG 638
Db 1851 AATTGATGAATGGTTATTAAAGAAATGCTAAAGAGAAAGCAATCAAGAAATTAAGAAAGC 1910
QY 639 AACTGCAATGTTAAATACACAGCTAAATATTAACTCTAAATCTACTAGTAAGAGTAATTT 698
Db 1911 AGAATCACTTCTGATTTATCTTCACTTAACTAAATCAATAAGAGCAAAACAGCTTGAAGCGT 1970
QY 699 AAGAGCTGCACTAGTAAGAGAGTTTACAAAATTTGAATGCTAGTTTCTTAATCTACAACTTT 758
Db 1971 AGAAGCAATTAAAGAACGAAATCTTAAAGCACACGCTGGAGAAAGAACACACAGAAATTA 2030
QY 759 AGCTGGTGATGACAGAAATACAAACAGCTATAGAGTAAGTAAGTAATTT--ACAATA 815
Db 2031 AGATGATATGCAACATATGAAGAGCAGAAAGCAGCAGCTTAAAGAAAGCTTTGAAATGA 2090
QY 816 TGATGGCGAGAAATCAGATCATTCAGCTGATGTTTAAAGAGAAATGTTTAAATGTTGTT 875
Db 2091 TGATGTTAAACAGCATACGAAATAGTTTCAAGTGCAGCGGAAGATCTACTATGTTATT 2150
QY 876 AGTAGGTGCAAAATGCACCTAGTAGATGATTTAGTTGCGGCTCTTTTAGCAGCAGAAAAAGA 935
Db 2151 AAGAGTTGAAGTTGCAGACGAGAAAGAACCCAGGTGAAGACACTCCAGAAAGTTCAAGAGG 2210
QY 936 TGCTCCACTATTATTAACTTCAAAAGATAAATTAGATTCGTGAGTAAATCTGAAATAA 995
Db 2211 TTACGCAACTTACCAAGAGCAGAAAGCAGCAGCTTAAAGAGCAATTAAGAAAGATAAAGT 2270
QY 996 GAGAGTTTATAGCTTAAATCTTCAACAGAAAGTAAACAGGAAACAGCTTATATAGCTGG 1055
Db 2271 TAACATGTCATACGAGTAGTTTCAAGTGCAGCGGAAGATCTACTATGTTTAAAT 2330
QY 1056 TGGAGTTAAATAGTGTATCTAAGAAAGTTGTAACAGAAATTAAGATCAATGGGATTAAAGT 1115
Db 2331 CGAGATTAAGAAAGATGAACACAGCTGGAAGAACCCAGGGAACCCAGGATCAAT 2390
QY 1116 TGAAGATCTCAGGTGATGATGATGAACTTTCTTTAAATATACAGGTGAAATAGG 1175
Db 2391 TGATGAATGGTTATTAAAGAAATGCTAAAGAGAGCGCAATCAAGAAATTAAGAAAGCAGG 2450
QY 1176 CTTAGATAATGATA 1189
Db 2451 AATCAGTTCTGACA 2464

RESULT 12
US-09-502-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-502-540-1357
Query Match 3.8%; Score 69.8; DB 4; Length 612;
Best Local Similarity 44.6%; Pred. No. 1.1e-05;
Matches 269; Conservative 0; Mismatches 334; Indels 0; Gaps 0;
QY 264 AGATAAAAAAGTTTATTCAGAACCAACTTTAACTACAGCTTAATGGAAATCAAGATTATGT 323
Db 6 AA 65
QY 324 AAGACAACTTTAAAAAAATTTAGATGCAGGAGAAATATGCTATTATATAGATTTAACCTTATA 383
Db 66 AA 125
QY 384 TAATGCTAAACCTGTTTGAATTTAAAGTAGTAGCAGCTAGTGAAAAACAGTAGTTGTATC 443
Db 126 AA 185
QY 444 TAGTGATGCGAAAAAATAGTGCAAAAGATATAGCTGAAAAATATGCTTTTGAACACAAAGA 503
Db 186 AA 245
QY 504 CTTAGAAAAATGCACCTAAAAAATATATAATGCTCAGATTTTCAGTAAAAAATCTAGTTACTA 563
Db 246 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 305
QY 564 TCAAGTAGTTCTTTTATCCAAAAGGAAAGAGATTTCAGAGGTTTCTCAACTTATAGAGCTAC 623
Db 306 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 365
QY 624 AATTTATATGAAGGAAGTGCATATGTAATACACAGTAAATATTAACCTCTAAATCTAC 683
Db 366 AA 425
QY 684 TAGTAAGAGTAATTTTAAAGACTGCAGTAGAAGAGTTTACAAAAATTCGAATCTAGTTATTC 743
Db 426 AAAAAAAAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 485
QY 744 TAATACTTACAACCTTTAGCTGTGATGACAGAAATCAAAACAGCTATAGAGATAGTAAAGA 803
Db 486 AA 545
QY 804 ATATTCAATAATGATGGCGAGAAATCAGATCAATTCAGCTGATGTTTAAAGAGATGTTAA 863
Db 546 AAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 605
QY 864 AAA 866
Db 606 AAA 608

RESULT 13
US-08-973-462-3
; Sequence 3, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3

```
; LENGTH: 1891
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1891)
US-08-973-462-3

Query Match      3.8%; Score 69.4; DB 3; Length 1891;
Best Local Similarity 42.4%; Pred. No. 1.9e-05;
Matches 495; Conservative 0; Mismatches 666; Indels 6; Gaps 2;

QY 664 ATATTAACTCTAAATCTACTAGTAGAGTAAATTTTAAAGACTGCGAGTAGAAGAGTTACAA 723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 ACATTAACTGAAAGTGTAGATGATAATAAAATTTTGAAGAAGCCGAAGATATAAAGGAA 61
QY 724 AAATTGAATGCTAGTTATTCTTAATCTACAACTTTTAGCTGGTGATCACAGAATACAAACA 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 AATATCTTTAAGTAATATAGAGAACCAGAAAGAAATATTATTGACAAATTTTAAAT 121
QY 784 GCTATAGAGATAAGTAAAGATATTACAATAATGATGGCGAGAAATCAGATCAATTCAGCT 843
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 AATATTGGACAAATTTCAAGAAAAACAAGAAAGTGATACAGAAAATGTACAAGTCAGTGAT 181
QY 844 GATGTTAAAGAGAAATCTTAAATATGTTGTATTAGTAGGTGCAAAATGCACCTAGTAGTGA 903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 GAACITTTTAAATGAATTTAATAATAGTGTAGATGTTAATGGAGAAGTAAAGAAAAATATT 241
QY 904 TTAGTTGCGGCTCTTTAGCAGCAGAAAAAGATGCTCCCACTATTATTAACTTTCAAAAGAT 963
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
242 TTGGAGAAAGTCAAGTTAATGACGATATTTTAAATAGTTTAGTAAAGATGTTCAACAA 301
QY 964 AAATTAGATCGCTGATAAAATCTGAAATTAAGAGAGTTTTAGACTTAAAAATCTCAACA 1023
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
302 GAACAACAACAAATGTTGAAGAAAAAGTTGAAGAAAGTGTAGA---AGAAATGACGAA 358
QY 1024 GAAGTAACAGGAAACACATTTATATAGCTGGTGGAGTTAATAGTCTATCTAAAGAGTT 1083
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 GAAAGTGTAGAAGAAATGTAGAAGAAATGTAGAAGAAATGACACGGAAGTGTAGCC 418
QY 1084 GTAAACAGAAATTGAATCAATGGGATTTAAAGTTTGAAGATTTCTCAGGTGATGATGATAT 1143
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 TCAAGTGTGGAAGAAAGTATAGCTTCAAGTGTGTGAAGATATAGATTTCAAGATTTGAA 478
QY 1144 GAAACTCTTTAAAAATAGCAGTGAAATAGGCTTAGATAATGATGAAGCTTATGATGTT 1203
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
479 GAAATGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTTGTAGCT 538
QY 1204 GGTGGAAACAGGATTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTACTAAATTAGAT 1263
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
539 CCAAGTGTGTAGAAGTGTGGCTCCCAAGTGTGGAAGAAAGTGTAGCTCCAAAGTTTGA 598
QY 1264 GGTAAATGGTGTGTAGATAGAACAAATGGACATGCTACTCCAAATGTTGTGTAGATGGA 1323
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
599 GAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCT 658
QY 1324 AAGCTGTATAAATATCTGATGACTTAGATAGTGTCTTAGGAAGCCGCTGATGTAGATATA 1383
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
659 CCAAGTGTGGAAGAAAGTGTAGCTGAAATGTTTGAAGAAAGTGTAGCTGAAATGTTGAA 718
QY 1384 ATAGGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATACAGTCTACTGCT 1443
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 GAAAGTGTAGCTGAAATGTTTGAAGAAAGTGTAGCTGAAATGTTTGAAGAAAGTGTAGCT 778
QY 1444 AAAGGCGTTTACAAGAGTTTAAAGCCGACGATAGACAAGACACTAACTCTGAAGTTATAAAA 1503
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
779 GAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTGTTGAA 838
QY 1504 ACATAATTGCTAATGATCTGAAATAGCTTAAAGCTGCGAGTTTATAGATAAAGATTCAGGT 1563
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
839 GAAATGTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAATTTGTAGTT 898
QY 1564 GCTTCAAGTAGTGCAGGAGTATTAAATTTCTATGTAGCTAAAGATGATGATCTACAAA 1623
```

```
Db 899 CCAAGTGTGGAAGAAAGTGTAGCTCCAAGTGTGTGAAGAAAGTGTAGCTGAAAATGTTGAA 958
QY 1624 GAAGATCAATTAGTTGATGCAATTAGCAGTAGAGAGCTG---TTGCTGGATATAAACTTGCT 1680
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
959 GAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCT 1018
QY 1681 CCAGTTGTTATAGCTACTGATTCITTTATCTTCTGATCAATCGTTCCTATTAAGCAAGTT 1740
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1019 GAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCGTAGCTCCAAAGTGTGAA 1078
QY 1741 GTAGGAGAAAAATATTCTTAAAGATTTTAAACAAGTTTGTCAAGGAATAGCTAAATTCAGTT 1800
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1079 GAAATCGTAGCTCCAACTGTTGAAGAAAGTGTGTCTGAAAAGCTTGCACAAATTTATCA 1138
QY 1801 ATAAACAAAAATGAAGATTTTATAGAT 1827
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1139 GACAATCTTTTAAAGTAAATTTATAGGT 1165

RESULT 14
US-09-141-047-7
; Sequence 7, Application US/09141047A
; Patent No. 6043085
; GENERAL INFORMATION:
; APPLICANT: Walker, David H.
; APPLICANT: Yu, Xue-Jie
; TITLE OF INVENTION: Ehrlichia canis 120-kDa Immunodominant Antigenic
; FILE OF INVENTION: Protein Gene
; FILE REFERENCE: D6143
; CURRENT APPLICATION NUMBER: US/09/141,047A
; CURRENT FILING DATE: 1998-08-27
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 7
; LENGTH: 2489
; TYPE: DNA
; ORGANISM: Ehrlichia canis
; FEATURE:
; LOCATION: -340..2149
; OTHER INFORMATION: Nucleotide sequence of gene encoding 120kDa
; OTHER INFORMATION: immunoreactive protein
US-09-141-047-7

Query Match      3.8%; Score 69.4; DB 3; Length 2489;
Best Local Similarity 42.2%; Pred. No. 1.9e-05;
Matches 583; Conservative 0; Mismatches 791; Indels 9; Gaps 3;

QY 287 AAACCTTTAACTACAGCTAATGGAATGAAGATTATGTAAAGACAACTTTAAAAAATTTAG 346
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
627 AAAAAGTATCTGAAACTAGTAAAGAGGAAGTACTCTCTGAAAGTTAAAGCAGAAGATTTGC 686
QY 347 ATGCAAGGAAATATGCTATTATAGATTTAACTTATAATAATGCTAAAACTGTTGAAATTA 406
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
687 AACCTGCTGTAGATGCTAGTATAGAACATTCATCAAGTGAAGTTGGAGAAAAAGTATCTA 746
QY 407 AAGTAGTAGCAGCTAGTGAAGAAAAACAGTAGTTGTATCTAGTGAATGCGAAAAATAGTCAA 466
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
747 AAACCTAGTAAAGAGGAAAGTACTCTCTGAAAGTTTAAAGCAGAAGATTTGCAACCTGCTGTAG 806
QY 467 AAGATATAGCTGAAAAATATGTTTGAAGACAAAGACTTAGAAAAATGCACATAAAAACTA 526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
807 ATGATAGTGTGGAAACATTCATCAAGTGAAGTTGGAGAAAAAGTATCTGAAACTAGTAAAG 866
QY 527 TAAATGCCCTCAGATTTCAAGTAAAACTGATAGTTACTATCAAGTAGTCTTCTTTATCCAAAG 586
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
867 AGGAAAAATACCTCTGAAGTTTAAAGCAGAAGATTTGCAACCTGCTGTAGTGTAGTATAG 926
QY 587 GAAAGAGATTACAAGGTTTCTCAACTTATAGAGTACAAATTTATTAATGAAGAACTGCAT 646
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
927 AACATTCATCAAGTGAAGTTGGAGAAAAAGTATCTTAAACCTAGTAAAGAAAGTACTTC 986
QY 647 ATGTGTAATACACACGCTAATATTAACTCTAAAAATCTACTAGTAAAGAGTAAATTTTAAAGACTG 706
```

```

Db      987 CTGAAGTTAAAGCAGAAAGATTTCGAACCTGCTGTAGATGATAGTGTGGAACATTCATCAA 1046
Qy      707 CAGTAGAAGAGTTACAAAATTCGAATGCTAGTTATTTCTATATCTACAACTTTAGCTGGTG 766
Db      1047 GTGAAGTTGGAGAAAAGATCTGAACCTAGTAAAGAAAGAAATATCTCTGAAGTTAAAG 1106
Qy      767 ATCAGAGATACAAACAGCTATAGAGATAGTAAAGAAATATTAACAATATGATGGCGAGA 826
Db      1107 CGGAAGATTGCAACCTGCTGTAGATGGTGTAGAACATTCATCAAGTGAAGTTGGGA 1166
Qy      827 AATCAGATCATTCAGCTGATGTTAAAGAGAAATGTTAAAGATGTTATGTTAGTAGGTGCAA 886
Db      1167 AAAAGATATCTGAACCTAGTAAAGAGGAAAGTACTCTGGAAGTTAAAGCAGAAAGATTGC 1226
Qy      887 ATGCACTAGTAGATGGAATGTTGGGCTCTTTAGCAGCAGAAAAGATGCTCCACTAT 946
Db      1227 AACCTGCTGTAGATG---ATAGTGTGGAACATTCATCAAGTGAAGTTGGGAGAAAAGTAT 1283
Qy      947 TATTAACTTCAAAGATAAATTAGATTCGTCAGTAAATCTGAAATCTGAAATAAAGAGAGTTTAG 1006
Db      1284 CTGAACCTAGTAAAGAGAAAATATCTCTGAAGTTAGCAGAAAGATTTGCCAACCTGCTG 1343
Qy      1007 ACTTAAAACTTCAAAGAGATTAACAGGAAAAACAGTTTATATAGCTGTGGGAGTTTAATA 1066
Db      1344 TAGATGCTAGTGTAGAACATTCATCAAGTGAAGTTGGAGAAAAGTATCTGAAACTAGTA 1403
Qy      1067 GTGTATCTAAAGAGTTGTACAGAAATTAAGATCAATGGGATTAAGAGTTGAAAGATTTCT 1126
Db      1404 AAGAGGAAAGTACTCTCTGAAGTTAAAGCAGAAAGATTTGCCAACCTGCTGTAGATAGTA 1463
Qy      1127 CAGGTGATGATAGATATGAACTCTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATG 1186
Db      1464 TAGAACATTCATCAAGTGAAGTTGGGAAAAGATATC---TGAAACTAGTAAAGAGGAAA 1520
Qy      1187 ATAAGGCTTATAGTGTGGGAAAGGATTTAGCAGATGCGCATGAGTATAGCTTTCAGTTG 1246
Db      1521 GTACTCTGAAGTTAAAGCAGAGATTTGCAACCTGCTGTAGATGGTGTAGAACATT 1580
Qy      1247 CTTCTACTAAATAGATGTAAGTGTGTGTAGATAGAACAAATGGACATGCTACTCCAA 1306
Db      1581 CATCAAGTGAAGTTGGGAAAAGTATCTGAAACTAGTAAAGAGGAAAATACTCTCTGAAG 1640
Qy      1307 TAGTTGTTGTAGATGGAAGAGCTGATAAATATCTGATGACTTAGATAGTT---TCTTAG 1363
Db      1641 TTAAGCAGAGATTTGCAACCTGCTGTAGATGTTAGTGTAGAACATTCATCAAGTGAAG 1700
Qy      1364 GAAGCGCTGATAGATATATAGTGGATTTGCAAGTGTATCTGAAAAGATGGAAGAAG 1423
Db      1701 TTGAGAAAAGTATCTGAACTAGTAAAGAGGAAATATCTCTGAAAGTTAAAGCGGAAG 1760
Qy      1424 CTATATCAGATGCTACTGTTAAAGCGGTTTAAAGAGTTAAAGGCGCAGTAGACAAGACA 1483
Db      1761 ATTTGCAACCTGCTGTAGATGTTAGTGTAGAACATTCATCAAGTGAAGTTGGAGAAAAG 1820
Qy      1484 CTAACCTGAAAGTTATAAACAATATATGCTATGATCTGAAATAGCTAAAGCTCAG 1543
Db      1821 TATCTGAACTAGTAAAGAGAAAGTATCTCTGAAAGTTAAAGCGGAGATTTGCAACCTG 1880
Qy      1544 TTTTAGATAAAGATTTCAAGTGTCTTCAAGTGTGATGAGGAGTATTTAAATTTCTATGTAG 1603
Db      1881 CTGTAGATGGTGTGGAACATTCATCAAGTGAAGTTGGGAGAAAAGTATCTGAGACTA 1940
Qy      1604 CTAAGAGTGGATCTACAAAAGAGATCAATTTAGTTGATGCAATTAGCAGTAGGAGCTGTG 1663
Db      1941 GTAAAGAGAAAGTACTCTCTGAAGTTAAAGCGGAAGATTTGCCAACCTGCTGTAGATGTA 2000
Qy      1664 CTG 1666
Db      2001 GTG 2003

```

RESULT 15
US-08-323-170B-1

```

; Sequence 1, Application US/08323170B
; Patent No. 5733772
; GENERAL INFORMATION:
; APPLICANT: Williamson, Kim C.
; TITLE OF INVENTION: Cloning and Expression of Plasmodium
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,170B
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,409
; FILING DATE: 29-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Quine, Jonathan A.
; REGISTRATION NUMBER: P-41,261
; REFERENCE/DOCKET NUMBER: 015280-113100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 149..9556
; US-08-323-170B-1

```

Query Match 3.8%; Score 69.4; DB 1; Length 9636;
Best Local Similarity 42.5%; Pred. No. 2.7e-05;
Matches 623; Conservative 0; Mismatches 821; Indels 21; Gaps 4;

```

Qy      370 GATTTAACTTATTAATAATGCTAAACCTGTTGAAATTAAGTAGTAGCAGCTAGTGAAGAAA 429
Db      569 GTTGTAGTAGTAGATGCAATTCGAATGTCAAGAAAGAAAATAATGATTAATAAGAAAT 628
Qy      430 ACAGTAGTGTCTATCTAGTGTGCAAAAATAGTGCAAAAGATATAGCTGAAAATATGTG 489
Db      629 ATAGATAAGAAATAAATAATGATATACAAATGTTGAGGAAAATAACATACAGGATACA 688
Qy      490 TTTGAAGACAAAGACTTAGAAAAATGCACATAAAACATATAAATGCCTCAGATTTTCAGTAAA 549
Db      689 TACGAAAATAAGAGTATGAAAGTGATGATACACTTATAGAAATGGTTTGATGATAATACA 748
Qy      550 ACTGATAGTTACTATCAAGTAGTCTTTATCCAAAGGAAAGAGATTTACAGGTTTCTCA 609
Db      749 AATGAAGAAAACCTTTTACTAACTTTTAAAAAGGCTGTGATGAAAATAATTTTCTTCA 808
Qy      610 ACTTATAGAGCTACAAATATATGAAGGAAGTGCATATGTTAATACACACAGTAATATTA 669
Db      809 CCCAAAGAAAACCTGTAGTACAAAAAACAATAAGTCTAATTTTATATAACAGT 868
Qy      670 ACTCTAAATCTACTAGTAAGAGTAATTTAAAGCTCAGCTAGAGAGTTA-----CAA 723
Db      869 TCGTTGAAATATATATATATGTTATTAACCCCTCGGATAGCTTTTAACTAGTAGTCGA 928

```

Qy	1798	GTTATAAACAAATGAAAGATTTAT	1822
D _b	2000	CCATTATAAAGTAAAAATAATAT	2024

Search completed: October 6, 2005, 10:25:02
Job time : 341 secs

Qy	724	AAATTGAATGCTAGTATTCTTAATPACAACTTTTAGTCTGGTGAATGACAGAAATACAAACA	783
Db	929	AACAGAAAATTTGGATGATGAGGAAGCATGTCGCCCAAGGATAAATTTTGTAAATAGATGATGAG	988
Qy	784	GCTATACAGATAAAGTAAAGAATATTACAAATAATGATGGCGAGAAATCAGATCAATTCAGCT	843
Db	989	GAAGACAGGAGNGGAAGAAAG	1048
Qy	844	GATGTTTAAAGAGAAATGTTTAAAAAATGTTTGTATTATTAGTAGGTCCAAATGCACCTAGTAGATGGA	903
Db	1049	GAGGAGGAAGAATATGATGATTATGTTTATGAAGAAAGTGGCGATGAAACAGAGAACAACAA	1108
Qy	904	TTAGTTGCGGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTTTAACTTCABAAGAT	963
Db	1109	TTACAGAGGAACAATCAGGAAGATAGGTGCTGTAATCTTCAGAGAAAGTTTTTAATGAT	1168
Qy	964	---AAATTTAGATTTCGTCACTAAAATCTCGAAATATAAGAGAGTTTTTAGACTTTAAAAACTTCA	1020
Db	1169	GAGGATCAAGATTCGTGAGAAGCACGGATGGAGATATGATAAGAGTTGACGAATATTAT	1228
Qy	1021	ACAGAAGTAAACAGAAAAACAGTTTTATATAGCTGGTGGAGTTAATAGTGTATCTTAAAGAA	1080
Db	1229	GAAGACCAAGATGCTGTACTCTTATGATAGTACAAATAAAAAATGAAGATGTAGATGAAGAG	1288
Qy	1081	GTTCCTAACAGAAATTTAGAAATCAATCGGATATAAAAGTTTGAAG---ATTCTCAGGTGATGAT	1137
Db	1289	GTAGGTCAAGAGGTAGTGTGAAGGTAGGTGTMAGAGGTAGTGMAGNGGTAGGTGAAGAG	1348
Qy	1138	AGATATCAAACTCTCTTTAAAAATAGCAGGTGAAATAGGCTTTAGATAATGATAAGGCTTAT	1197
Db	1349	GTAGGTCAAGAGGTAGCTGAAGAGGTAGGTGAAGAGGTAGGTGCAAGAAGAGGTGAAGAG	1408
Qy	1198	GTAGTTGGTGGACAGAGATTTAGCAGATGCCATGAGTATAGCTTCAGTTGCTTCTACTAAA	1257
Db	1409	GTAGGTCAAGGGGTAGGTGTAAGAGGTAGGTGAAGAAGAGGTGAAGAGGTAGGTGAAGAA	1468
Qy	1258	TTAGATGATTAATGGTGTGTAGATAGAAACAAATGGCATGCTACTCTCAAATAGTTGTGTGA	1317
Db	1469	GAAGGTCAATATGTPAGATGAAGAAGAAGCAAGGTGCAAAATATATCCATTTGGTGATGA	1528
Qy	1318	GATGGAAAAGCTGATAAAATATCTGATGACTCTGATAGATTTCTTAGGAAGCGCTGATGTA	1377
Db	1529	GAAGAAAAAGATGA-----AGGTGGAGAAGTTTTTACCTATGAAAAGACCGAGGTT	1579
Qy	1378	GATATAATAGTGTGATTTGCAAGTGTTATCTGAAAAGATGGAAGAAGCTATATCAGATGCT	1437
Db	1580	GATAAAACAGATTTGTTTAAATTTATAGAGGGGGTCAAGGAGATGATGATATATAAGTG	1639
Qy	1438	ACTGGTAAAGCGGTTACAAGAGTTAAAGCGGACGATAGACAAGACACTAACTCTGAAGTT	1497
Db	1640	GATGGTTCCAAAGTTTTTATTAGATGATGATACAAATAGTAGAGTATCTTAAAAACATCT	1699
Qy	1498	ATAAAAACATATTATGCTTAATGATCTAGAAATAGCTTAAAGCTGCAGCTTTTAGTATAAAGAT	1557
Db	1700	GCACGAGATGGTGAATATGTTGTAATATGGTGAAGCTGTCTGAAAGATGAGAAAAATGTTATA	1759
Qy	1558	TCAGGTGCTTCAAGTAGTGATGACGAGGATTTTAAATTTCTATGTAGCTAAAGATGGATCT	1617
Db	1760	AAAATAATTAGAGTGTGTTTACAAGTGGTGCAATTACCAAGTGTAGGTGTTGATGAGTTA	1819
Qy	1618	ACAAAAGAGATCAATTTAGTTGATGCAATTAGCAGTAGGAGCTGTTGCTGGATATAAACTTT	1677
Db	1820	GATAAAAATCGATTTGTCTATATGAAACAACAGAAAGTGGAGATATCTGCTGTATCCGAAGAT	1879
Qy	1678	GCTCCAGTTGTATTAGTACTGATTTCTTTATCTTCTGATCAATCGGTTGCTATAAGCAAA	1737
Db	1880	TCATATGATAATAATGCAATCTAATAATACAAATAAGAAATACGTTTGTGATTTTACAGAT	1939
Qy	1738	GTTCGTAGGAGAAAAATATCTTAAAGATTTTAAACAAGATTTGGTCAAGGAATAGCTAAATCA	1797
Db	1940	CAATTTAAAAACCAACAGAAAGTGGTCTTAAAGTAAAAAAATGTCAAGTAAAAAGTTAATGAG	1999

Result No.	Score	Query Match	Length	DB	ID	Description	
C	1	113.8	6.2	1805	9	CL080711	CH216-159
	2	111.6	6.1	1784	9	CL081992	CH216-165
	3	106.2	5.8	2001	9	CL118787	ISB1-72M2
	4	104.6	5.7	1416	9	CL499036	SAIL 662_
	5	104.2	5.7	1843	9	AG435185	Mus muscu
	6	104	5.7	1378	9	AG350209	Mus muscu
C	7	103	5.6	1436	9	CL057761	CH216-860
	8	102	5.6	1632	9	CL082569	CH216-167
C	9	101.8	5.6	1376	9	CG747831	PO41-3-B0
	10	101.8	5.6	1493	9	CL078589	CH216-151
C	11	101.8	5.6	1506	9	AG278469	Mus muscu
	12	101.8	5.6	1608	9	CL118721	ISB1-72J8
	13	101.6	5.6	1594	9	CL078613	CH216-151
	14	100.4	5.5	1162	9	CL077122	CH216-143
15	100.4	5.5	1519	9	AG386893	Mus muscu	
16	100.4	5.5	1531	9	CG748014	PO41-4-B0	
C	17	100.4	5.5	1542	9	AG386981	Mus muscu
C	18	100.4	5.5	1566	9	CG757757	PO53-1-D0
C	19	100.2	5.5	1217	9	CL078190	CH216-148
	20	100.2	5.5	1599	9	CL083840	ISB1-2H14
	21	99.8	5.5	1594	9	CL038406	CH216-46A
22	99.6	5.4	1842	9	CL044158	CH216-59P	
23	99.4	5.4	1367	9	CL082650	CH216-168	
24	99.4	5.4	1434	9	AJ592058	Arabidops	

```
Db 1215 TAATAATAAAAAATATTGATTAAAAATAAAAAAATATATATTAATAAATGAAAAATG 1156
Qy 146 ACTTAGTAAAGGATATTTTAGCGACGCAAAACCTTAAACAAGTGCGAGTTATTTTGAACA 205
Db 1155 AAAAAAATAAAAAATATAAAAAAATAAAAAATAAAATATATATAAACAATGATAATGATAA 1096
Qy 206 AGATACAAAGTTACTTCTATGATGCAAAACAGAAAGATTCTTCAACTCCACTGGAG 265
Db 1095 AATATATTAATTATTAATAAATGAAAAAATAAAAAAATTTAAAAAATAAATAAAAAA 1036
Qy 266 ATAAAAAGTTTATTTCAGAACAACT - -TTAACTACAGCTAAATGGAAATGAAGATTATGT 323
Db 1035 AAAAAAGAAATATTTTATTAATAAATAATTAATATAGATATTAATAAAAAAATAAATA 976
Qy 324 AAGACAACTTTAAAAAATTTAGATGCGAGAAATATGCTATTATATAGATTAACTTTATAA 383
Db 975 AAAAAAAGAAAAAATTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 917
Qy 384 TAATGCTAAACCTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAAAAAACAAGTAGTTGATC 443
Db 916 AAAAAAATAAAAAAATTTAAGAACAAAAATAAAAAAATAAAAAAATAAAAAAATAAAA 857
Qy 444 TAGTGATGCAAAAAATAGTGCAAAAAGATATAGCTGAAAAATATGCTTTTGAAGACAAAGA 503
Db 856 AATATATGAATAAAAAA - - - - -AAAAATTAATAAAAAAATAAAGAAAAAATAAAAAA 801
Qy 504 CTTAGAAATGCACTAAAAATATATAATGCCTCAGATTTTCAGTAAAACTGATAGTTACTA 563
Db 800 AAAAAAATATATAGATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 741
Qy 564 TCAGTAGTCTTTATTCACAAAGGAAAGAGATTAACAAGTTTCTCAACTTATAGAGCTAC 623
Db 740 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 681
Qy 624 AATATTATATGAGGAACGTCATATGTTGTAATACACAGTAATATTAACCTTAAATCTAC 683
Db 680 AATAAATAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 621
Qy 684 TAGTAAGAGTAATTTAAAGACTGCGAGTAAGAGTTTACAAAAATTTGAATCTGATTTATTC 743
Db 620 TAAAGTGAAATAAATAATATAAATGTAATAATGTAATAATGTAATAATGTAATAATGTA 561
Qy 744 TAATACTACAACCTTTAGCTGGTGACAGATACAAACAGCTATAGAGATAGTAAGA 803
Db 560 AATAAAAAATAAAAAATTTATATAAAAAATAAAAAAATANAGAAAAATAAAAAATAAATTA 501
Qy 804 ATATTACAATATGATGGCGAGAAATCAGATCAATCAGCTGATGTTTAAAGAGAAATGTTAA 863
Db 500 ATTATAAAAAATAATCAATAATAAATAAATAATAAATAAATAAATAAATAAATAAATAA 441
Qy 864 AATGTTGTTATTAGTAGTGCAAAATGCATAGTAGATGATTAGTTGCGGCTCCTTTAGC 923
Db 440 TAATAATTATATAAAATGATAAAATGAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 381
Qy 924 AGCAGAAAAAGATGCTCCACTATTATTAACCTTCAAAAGATAAATTAGATTCTGCAGTAA 983
Db 380 AAAAAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 321
Qy 984 ATCTGAAATAAAGAGAGTTTGTAGCTTAAAAACTTCAACAGAAAGTAACAGGAAAAACAGT 1043
Db 320 AATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 261
Qy 1044 TTATATAGCTGGTGGAGTTTAATAGTGATCTAAGAAAGTTGTAAACAGAAATTAGATCAAT 1103
Db 260 TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 201
Qy 1104 GGGATTTAAAGTTGAAGAGTTCTCAGGTGATGATAGATGAACCTTCTTTAAAAATAGC 1163
Db 200 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 141
Qy 1164 AGGTGAAATAGGCTTAGATAATGATAAGGCTTATGTAGTTGGTGGAAACAGGATTAGCAGA 1223
Db 140 AATTNANNANCAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 81
```

```
Qy 1224 TCCCATAGATATAGCTTCAGTTGCTTCTACTAAATTAGATGTAATGTTGTGTAGATAG 1283
Db 80 NNTAAAAAGATATTAAANNCCCTTTGAAAAAAGAAAAAGATAATTCGNATNAAANCT 21
Qy 1284 AACAAATGGACA 1295
Db 20 AAAAAATTTAAAA 9

RESULT 2
CL081992
LOCUS
DEFINITION
  CH216-165M9_RNA.1 CH216 Xenopus tropicalis genomic clone
  CH216-165M9, genomic survey sequence.
ACCESSION
  CL081992
VERSION
  CL081992.1 GI:40537905
KEYWORDS
  GSS.
ORGANISM
  Xenopus tropicalis (western clawed frog)
  Xenopus tropicalis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
  Xenopodinae; Xenopus; Silurana.
  1 (bases 1 to 1784)
REFERENCE
  1 Krenitski,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
  Mardis,E. and Wilson,R.
  A physical map of the xenopus tropicalis genome
  Unpublished (2003)
  Contact: Richard K Wilson
  Genome Sequencing Center
  Washington University School of Medicine
  Email: submissions@watson.wustl.edu
  Insert Length: 175000 Std Error: 0.00
  Seq primer: RM4 ctcaagggcacgcgctcgagc
  Class: BAC ends
  High quality sequence start: 409
  High quality sequence stop: 503.
  Location/Qualifiers
    1..1784
      /organism="Xenopus tropicalis"
      /mol_type="genomic DNA"
      /strain="Nigerian frog"
      /db_xref="taxon:8364"
      /clone="CH216-165M9"
      /sex="male"
      /cell_line="Stock 248 F7A2, inbred N7"
      /clone_lib="CH216"
      /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
      BAC library"

FEATURES
  source
    1..1784
      /organism="Xenopus tropicalis"
      /mol_type="genomic DNA"
      /strain="Nigerian frog"
      /db_xref="taxon:8364"
      /clone="CH216-165M9"
      /sex="male"
      /cell_line="Stock 248 F7A2, inbred N7"
      /clone_lib="CH216"
      /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
      BAC library"

ORIGIN
  Query Match 6.1%; Score 111.6; DB 9; Length 1784;
  Best Local Similarity 45.0%; Pred. No. 2.9e-12;
  Matches 497; Conservative 0; Mismatches 604; Indels 3; Gaps 2;
  Qy 87 AATATCACTACAAGATGGTACAATAAGATAAGTATACAGTATCAATACATACTAAAGCTAGTGA 146
  Db 480 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAAAAAAAAAAAAAAAA 539
  Qy 147 CTTAGTAAGGATATTTTAGCAGCACAAAACCTTACACAGCTGCAGTATTTTTGAACAA 206
  Db 540 AAAAAAAAAAAAAAAAAAGAAAGAAAAAAAAAAAAAAAAAGAAAGAAAGAGAGAAAAA 599
  Qy 207 AGATACAAAAAGTTACTTTCTATGATGCAAAATGAGAAAGATTTCTTCAACTCGGAGA 266
  Db 600 AAAAAAAAAATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 659
  Qy 267 TAAAAAAGTTTATTCAGAACAACTTTAACTACAGCTAATGGAAATGAAGATTATGTAAA 326
  Db 660 AATAATATAAAAAAAAAAAAAAAAAAGAAATATATATAAATAAATAAATAAATAAATAA 719
  Qy 327 GACAACTTTAAAAATTTAGATGCGAGGAGATATGCTATTATAGATTAACTTTAATAATA 386
```



```
Db 759 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATA 818
QY 807 TTACAAATATGATGCGGAAATCAGATCATTCAGCTGATGTTAAAGAGAATGTTAAAAA 866
Db 819 AATAAATAAGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 878
QY 867 TGTGTATTAGTAGTGCAATGCACTAGTAGATGGATTAGTTGGCGGCTCCTTTAGCAGC 926
Db 879 AGAAAAAAGAAAAAGAAAAAAGAAAAAG--AATAAATAAATAAATAAATAAATAAATAAATAA 936
QY 927 AGAAAAAGATGCTCCACTATTATTAACTTCAAAAGATAAATTAGATTCCTCAGTAAATC 986
Db 937 AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAATAAGAAAAAACAATAATAAAAA 996
QY 987 TGAATAAAGAGAGTTTGTAGCTTAAACCTTCAACAGAGTAAACAGGAAAAACAGTTTA 1046
Db 997 AGAAAAAAGAAAAAATAATAAATAAATAAAGAAAAAATAAATAAATAAAGAAATAGAGAAAA 1056
QY 1047 TATAGCTGGTGAGTTTAAATAGTGTATCTAAAGAAGTTGTAACAGAAATTAGATCAATGGG 1106
Db 1057 AATAAACAACAAATAGAAAAAACAAGAAAAAATAAATAAATAAATAAATAAAGAAAAA 1116
QY 1107 ATTAAAAAGTTGAAGATTCTCAGGTGATGATAGATATGAACCTTCTTTAAAAAATAGCAGG 1166
Db 1117 AAAAGAGATAAATAAATAAATAAGAGAAAAAAGAAAAAAGAAAAAATAAAGTATA 1176
QY 1167 TGAAA 1171
Db 1177 AAAAA 1181
```

```
RESULT 4
LOCUS CL499036
DEFINITION SAIL 662 F12.v3 SAIL Collection Arabidopsis thaliana linear GSS 01-APR-2004
SOURCE SAIL_662_F12.v3, genomic survey sequence.
ACCESSION CL499036
VERSION CL499036.1 GI:45995368
KEYWORDS GSS.
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1416)
Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmery,B.,
Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
22356987
12468722
Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS828833; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1. .1416
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL 662 F12.v3"
/clone_lib="SAIL_Collection"
```

FEATURES
source

/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN	Query Match	Score	104.6;	DB	9;	Length	1416;			
	Best Local Similarity	40.3%;	Pred. No.	7.5e-11;						
	Matches	438;	Conservative	0;	Mismatches	644;	Indels	5;	Gaps	2;
QY	87	AATATCACTACAGATGGTACAAATGATAAGTATACAGTATCAATATCTAAATCTAAAGCTAGTGA	146							
Db	303	ANNNNNNNNNATTTTTTTTTTATANAATAAATAAATAAATAAATAAATAAATAAATAAATAA	362							
QY	147	CTTAGTAAGGATATTTTTAGCAGCACAAAACCTTAACAACAGGTGCAGTTATTTTGAACAA	206							
Db	363	ANAGAAATATATATNTNAAAAAAGAAAAAGAAAAAATAAATAAATAAATAAATAAATAA	422							
QY	207	AGATACAAAAGTTTACTTTCTATGATGCAATGAGAAAGATTCTTCAACTCCAACCTGGAGA	266							
Db	423	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	482							
QY	267	TAAAAAGTTTATTCAGAACAACTTTTAACCTACAGCTAATGGAATGAAGATTATGTAA	326							
Db	483	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	542							
QY	327	GACAACTTTTAAAAATTTAGATGCAGGAGAATATGCTATTATAGATTAACTTATAATAA	386							
Db	543	NNNNNNNNAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	602							
QY	387	TGCTAAAACTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAAAAAACAAGTAGTTGTCTAG	446							
Db	603	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	662							
QY	447	TGATCGCAAAATAGTGCAAAAGATATAGCTGAAAAATATGTTGTTGAAGACAAAGACTT	506							
Db	663	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	722							
QY	507	AGAAATGCACCTAAAAACTATAAATGCCTCAGATTTTCAGTAAAACTGATAGTTACTATCA	566							
Db	723	AAAGAAAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	782							
QY	567	AGTAGTCTTTATCCAAAAGGAAAGAGATTACAAGTTTCTCAACTTATAGAGCTACAAA	626							
Db	783	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	842							
QY	627	TATATGAAGGAACCTGCATATGTTATACACCAGTATATTAATCTCTAAATCTACTAG	686							
Db	843	TAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	902							
QY	687	TAAGAGTAAATTAAAGACTGCAGTAGAAGAGTTACAAAAATTTGAATGCTAGTTATCTAA	746							
Db	903	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	959							
QY	747	TACTACAACTTTAGCTGGTGTATGACAGAAATCAACACAGCTATAGAGATAAGTAAAGATA	806							
Db	960	NNAAAAAAGAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1019							
QY	807	TTACATATATGATGCGGAGAAATCGAGTCACTTCAGCTGATGTTAAGAGAAATGTTAAAA	866							
Db	1020	NATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1079							
QY	867	TGTTGTTATTAGTAGTGCAATGCACTAGTAGATGATTTAGTTGCGGCTCCTTTAGCAGC	926							
Db	1080	NAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1139							
QY	927	AGAAAAAGATGCTCCACTATTATTAACTTCAAAAGATAAATTAGATTTCGTCAAGTAAATC	986							
Db	1140	NNATAAATANGNNNNNNNAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1197							
QY	987	TGAATAAAGAGAGTTTGTAGCTTAAACCTTCAACAGAGTAAACAGGAAAAACAGTTTA	1046							
Db	1198	NNAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1257							
QY	1047	TATAGCTGGTGAGTTTAAATAGTGTATCTAAAGAAGTTGTAACAGAAATTAGAATCAATGGG	1106							


```

sequence.
ACCESSION AG350209
VERSION AG350209.1 GI:47923519
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus

REFERENCE
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE BAC end Sequences of Library MSMg01
JOURNAL Unpublished
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
Library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACE3.6
Vector : EcoRI
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
1. .1378
Location/Qualifiers
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-146M02.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 5.7%; Score 104; DB 9; Length 1378;
Best Local Similarity 45.9%; Pred. No. 9.9e-11;
Matches 430; Conservative 0; Mismatches 501; Indels 6; Gaps 2;

QY 108 AAATGATAGTACAGTATCAAACTAAAGCTAGTACTAGTAAAGGATATTTAGC 167
DB 1346 ATAATAATAATTAATAATATATCTATATAAAAAATAATAATAATTAATAATTAATACT 1287
QY 168 AGCAGAAAACCTTAACAACAGGTGCAGTTATTTTGAACAAGATACAAAA--GTTACTTT 224
DB 1286 ATAATAATAATTAATAATTAATAATAACAATAATTAATAACAATACTAAATAA 1227
QY 225 CTATGATGCAATGAGAAAGATTCTTCAACTCCAAGTGGAGATAAAAAAGTTTATTCAGA 284
DB 1226 ATATATATTAAATATATAAAAAATAATAATAATAATAATAATAATAATAATAATAA 1167
QY 285 ACAAACTTTAATACAGCTAATGGAATGAAGA---TTATGTGAAGACAACTTTAAAAA 341
DB 1166 ATAAAAATTTAAATATATAAAAAATAATAAAAAATAATAAAAAATAAAAAATAATAAT 1107
QY 342 TTATGATGAGGAGATATGCTATTATAGATTAACTTATATAATGCTAAAACTGTTGA 401
DB 1106 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1047
QY 402 AATTAAGTAGTACGAGCTAGTGAAAAAACAAGTAGTTGATCTAGTGATCGGAAAAATAG 461
DB 1046 AAAAAATTTAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAA 987

```

```

QY 462 TGCRAAGATATAGCTGAAABAATATGTGTTTGAACACAAAGACTTTAGAAAAATGCACATAA 521
DB 986 TAATAAAAAAATAATAAAAAATAATAATAAAAAACAAAAATAATAATAATAATAAAAAAAA 927
QY 522 AACTATAAATGCCTCAGATTTTCAGTAAACCTGATAGTTACTATCAAGTAGTTCTTTATCC 581
DB 926 AAATAACACATAATTTAAATATAATAAAAAATAAAAAATAAAAAATAATAATAATAATA 867
QY 582 AAAAGGAAGAGATTACAAGGTTTCTCACTTATAGAGCTACAATTTATTAATGAAGGAAC 641
DB 866 TAAAAAATAAATAATAAAAAATAATAAAAAATAATAAAAAATAATAAAAAATAATAATA 807
QY 642 TGCATATGGTATACACCAGTAATAATTAACTCTAAATCTACTACTAGAGAGTAATTTAAA 701
DB 806 AAATAAATAAATAAATAAAAAATAAAAAATAATAATAAAAAATAATAATAATAATAATA 747
QY 702 GACTGCAGTAGAAGAGTTTACAAAAATTCGAATGCTAGTTATTCTTAATCTCAACACTTTAGC 761
DB 746 TTATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 687
QY 762 TGGTGATGACAGAAATACAAAAGCTATAGAGATAAGTAAGAATATTACAATAATGATGG 821
DB 686 AATAAATAAATAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAATAA 627
QY 822 CGAGAAATCAGATCAATTCAGCTGATGTTAAAGAGAGATCTTAAAGATGTTGTTATTAGTAGG 881
DB 626 AAAAAAATAAATAATAATAAAAAATAAAAAATAAAAAATAAAAAATAATAATAATAA 567
QY 882 TGCRAATGCACCTAGTAGATGATTTGCGGCTCTTTTAGCAGCAGAAAAAGATGCTCC 941
DB 566 ATATAAATAAATAAATAATAATAATAATAAATAAATAAATAAATAAATAAATAAATAA 507
QY 942 ACTATTATTAACTTCAAAAGATAAATTAGATTGCTGAGTAAATCTGAAATAAGAGAGT 1001
DB 506 AATAAAAAATAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAGGA 447
QY 1002 TTTAGACTTAAAACTTCAACAGAGAGTAACAGGAAAA 1038
DB 446 GTAATTATTATAATAATAATAAATAAAAAATAAATAAATAAATAAATAAATAAATA 410

RESULT 7
LOCUS CL057761
DEFINITION CH216-8609 RMI.1 CH216 xenopus tropicalis genomic clone CH216-8609,
genomic survey sequence.
ACCESSION CL057761
VERSION CL057761.1 GI:40513674
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1436)
AUTHORS Krenitski,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM1 TACGACTCACTACTAGGGAGA
Class: BAC ends
High quality sequence start: 220
High quality sequence stop: 1095.
Location/Qualifiers
1. .1436
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"

FEATURES
source

```

/db_xref="taxon:8364"
/clone="CH216-8609"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: PTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

Query Match 5.6%; Score 103; DB 9; Length 1436;
Best Local Similarity 43.4%; Pred. No. 1.6e-10;
Matches 472; Conservative 0; Mismatches 615; Indels 0; Gaps 0;

```
QY 87 AATATCACTACAGAGTGGTCAAAATGATAGTATACAGATGATCAAACTAAAGCTAGTGA 146
DB 329 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 388

QY 147 CTTAGTAAGGATATTTTAGCAGCACAAAACCTTAAACACAGGTGCAGTTATTTTGAACAA 206
DB 389 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 448

QY 207 AGATACAAAGTTACTTCTATGATGCAATGAGAAAGATCTTCAACTCCACTGGAGA 266
DB 449 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 508

QY 267 TAAAAAGTTTATTCAGAACAACTTTAACTACAGCTTAATGGAATGAGATTTATGAAA 326
DB 509 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 568

QY 327 GACAACTTTAAAAATTTAGATCAGGAGAAATATGCTATTATATAGATTTAACTTATATAA 386
DB 569 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 628

QY 387 TGTCTAAACTGTTGAATTAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTTGTATCTAG 446
DB 629 AAGAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 688

QY 447 TGATGCGAAAAATAGTGCRAAGATATAGCTGAAAAATATGTCTTGAACACAAAGACTT 506
DB 689 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 748

QY 507 AGAAATGCACTAAAACTAAATGCTCAGATTTTCAGTAAACTGATAGTTACTATCA 566
DB 749 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAA 808

QY 567 AGTAGTTCTTTATCCAAAGGAAGAGATTACAGGTTCTCAACTTATAGAGCTACAA 626
DB 809 AAAAAAGACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 868

QY 627 TTATATGAGGAACTGCATATGTTATACACAGTAAATATTAACCTTAAATCTACTAG 686
DB 869 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 928

QY 687 TAAGAGTAATTTAAAGCTGCAGTAGAAGAGTTACAAAAATGGAATCTAGTTATTCTAA 746
DB 929 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 988

QY 747 TACTACAACTTTAGCTGGTGAATGACAGAAATACAAACAGCTATAGAGATAAGTAAGATA 806
DB 989 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1048

QY 807 TTCAATAATGATGCGCGAGAAATCAGATCATTCAGCTGATGTTTAAAGAGAAATGTTAAAA 866
DB 1049 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1108

QY 867 TGTGTATTAGTAGTGCATAGTAGAGTAGGATTAGTTAGTCGCGCTCTCTTAGCAGC 926
DB 1109 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1168

QY 927 AGAAAAAGATGCTCCACTATTATTAACTTCAAGATATAATTTAGATTGCTCAGTAAATC 986
DB 1169 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1228
```

```
QY 987 TGAATAAAGAGAGAGTTTTAGACTTTAAACCTTCAACAGAGTAAACAGGAAAAACAGTTTA 1046
DB 1229 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1288

QY 1047 TATAGCTGGTGGAGTTAATAGTGTATCTAAAGAACTTGTAAACAGAAATTAGATCAATGG 1106
DB 1289 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1348

QY 1107 ATTAAAAAGTTGAAAGATTCTCAGGTGATCATGATATGAAACTCTTTTAAAAATAGCAGG 1166
DB 1349 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1408

QY 1167 TGAATA 1173
DB 1409 AGAAAAA 1415
```

RESULT 8
LOCUS CL082569 1632 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-167P5 Sp6.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CH216-167P5_ genomic survey sequence.
VERSION CL082569
KEYWORDS CL082569.1 GI:40538482
SOURCE GSS.

ORGANISM Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 1632)
AUTHORS Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.

TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu

Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 1062
High quality sequence stop: 1114.

FEATURES Location/Qualifiers

source 1..1632
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-167P5"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: PTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN

Query Match 5.6%; Score 102; DB 9; Length 1632;
Best Local Similarity 42.7%; Pred. No. 2.6e-10;
Matches 471; Conservative 0; Mismatches 633; Indels 0; Gaps 0;

```
QY 87 AATATCACTACAGAGTGGTACAAATGATAGTATACAGATGATCAAACTAAAGCTAGTGA 146
DB 430 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 489

QY 147 CTTAGTAAGGATATTTTAGCAGCACAAAACCTTAAACACAGGTGCAGTTATTTTGAACAA 206
DB 490 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 549

QY 207 AGATACAAAGTTACTTCTTATGATGCAATGAGAAAGATCTTCAACTCCACTGGAGA 266
DB 550 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 609
```


QY 627 TTATAATGAAGCACTGCATATGGTAATACACAGTAATATTAACTCTAAATCTACTAG 686
Db 632 AA 573
QY 687 TAAGAGTAATTTAAAGCTGCTAGAGAGTTTACAAAATTGAATGCTAGTTTCTTAA 746
Db 572 AA 513
QY 747 TACTACAACTTTAGCTGGTGCATGACAGATACAAACAGCTATAGAGTAAGTAAAGCAATA 806
Db 512 ANAAA 453
QY 807 TTAACAATAATGATGGCGAGAAATCAGATCACTTCAGCTGTGTTTAAAGAGAAATTTAAAAA 866
Db 452 AA 393
QY 867 TGTGTATTAGTAGGTGCAATGCATGATAGATGGAATTAGTTGGCGCTCTTTAGCAGC 926
Db 392 AA 333
QY 927 AGAAAAAGATGCTCCACTATTATTAACTTCAAAAGATTAATTTAGATTTCGTAGTAAATC 986
Db 332 AA 273
QY 987 TGAATAAAGAGAGTTTTCAGCTTAAAACTTCAACAGAAAGTAAACAGGAAAAACA 1041
Db 272 AA 218

RESULT 10
LOCUS CL078589 1493 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-151K5_Sp5.1 CH216 Xenopus tropicalis genomic clone
CH216-151K5, genomic survey sequence.

ACCESSION
VERSION CL078589
KEYWORDS CL078589.1 GI:40534502
SOURCE GSS.

ORGANISM Xenopus tropicalis (western clawed frog)

Xenopus tropicalis
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 1493)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E. and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 175000 Std Error: 0.00

Seq primer: Sp5 atctgcgcttcgatact

Class: BAC ends

High quality sequence start: 1009

High quality sequence stop: 1070.

Location/Qualifiers

1..1493

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/strain="Nigerian frog"

/db_xref="taxon:8364"

/clone="CH216-151K5"

/sex="male"

/cell_line="Stock 248 F7A2, inbred N7"

/clone_lib="CH216"

/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis

BAC library"

ORIGIN

Query Match 5.6%; Score 101.8; DB 9; Length 1493;
Best Local Similarity 42.4%; Pred. No. 2.8e-10;
Matches 464; Conservative 0; Mismatches 631; Indels 0; Gaps 0;

QY 96 ACAGATGCTACAAATGATTAAGTATACAGTATCAATACTAAAGCTAGTACTAGTAAA 155
Db 373 AA 432
QY 156 GGATATTTTTCAGCAGCAAACTTTAAACACAGCTGCAGTTATTTTGAACAAGATACAAA 215
Db 433 AA 492
QY 216 AGTTACTTTCTATGATGCAATGAGAAAGATTCTTCAACTCCAACTGGAGATAAAAAAGT 275
Db 493 ANAAA 552
QY 276 TTATTCAGAACAACTTTTAACTACAGCTAATCGAAATGAAGATTATGTTAAGCAACTTT 335
Db 553 AA 612
QY 336 AAAAAATTTAGATGCGAGGAGATATGCTATTATAGTTTAACTTATATAATGCTAAAC 395
Db 613 NAAA 672
QY 396 TGTGTAAATTAAGTAGTAGCAGCTAGTCGAAAAACAGTAGTTGTATCTAGTATGCGAA 455
Db 673 AA 732
QY 456 AAATAGTGCAGAAAGATATAGCTGAAAAATATGTTTGAAGACAAAGACTTAGAAATGC 515
Db 733 AA 792
QY 516 ACTAAAACTATAAATGCTCAGATTTTCAGTAAACTGATAGTTACTATCAAGTAGTTCT 575
Db 793 AA 852
QY 576 TTATCCAAAGGAGAGAGATTACAAGTTTCTCAACTTATAGAGCTACAAATTTATATGA 635
Db 853 AA 912
QY 636 AGGAATCGCATATGTTAATACACAGCTAATATTAATCTTAAATCTACTAGTAGAGTAA 695
Db 913 AA 972
QY 696 TTTTAAAGACTGCTAGTAGAGAGAGTTACAAAATTTGAATGCTAGTTTCTTAATCTACAC 755
Db 973 TAAA 1032
QY 756 TTTTACCTGCTGATGACAGAGATACAAACAGCTATAGAGATAAGTAAAGATTTTACATAA 815
Db 1033 AA 1092
QY 816 TGATGCGAGAAATCAGATTCATTCAGCTGATGTTTAAAGAGAAATGTTTAAATTTGTATT 875
Db 1093 AA 1152
QY 876 AGTAGTGCAATGCTAGTAGATGGAATGATGTCGGGCTCTTTAGCAGCAGAAAAGA 935
Db 1153 ATAAA 1212
QY 936 TGTCTCACTATTATTAACTTCAAAAGATTAATTTAGATTTCGTAGTAAATCTGAAATAAA 995
Db 1213 AA 1272
QY 996 GAGAGTTTTTACAGCTTAAAAAATCTTCAACAGAGTAAACAGGAAAAACAGTTTATATAGCTGG 1055
Db 1273 AA 1332
QY 1056 TGGAGTTTAAATGATGTTATCTAAAGAGTTGTAACAGAAATTAAGATTCATGGATTAAAGT 1115
Db 1333 AAAATRAAAATTAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTAA 1392
QY 1116 TGAAGATTTCTCAGGTGATGATAGATATGAACTTCTTTAAATATACGAGTGAATAGG 1175
Db 1393 AAATAAATAATAAANATATAAAAAAAAAAAAAAAAAAAAAAAAAAATTTAA 1452

```

QY 1176 CTTAGATATGATAA 1190
DB 1453 NTAATAATAATAAA 1467

RESULT 11
AG278469/c
LOCUS Mus musculus molossinus DNA linear GSS 02-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-048M18.T7, genomic survey
sequence.
ACCESSION AG278469
VERSION AG278469.1 GI:47851346
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1506)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Mashira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-mi-ku, Yokohama, Kanagawa 230-0045, Japan
[E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170]
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY : pBAC3.6
Vector : EcoRI
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
1. .1506
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-048M18.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 5.6%; Score 101.8; DB 9; Length 1506;
Best Local Similarity 44.7%; Pred. No. 2.8e-10;
Matches 489; Conservative 0; Mismatches 597; Indels 9; Gaps 2;

QY 80 CAGATGTAATCACTACAGAGTGGTACAAATGATAGTATACAGTACCAATCTAAAG 139
DB 1244 CAAAAAATAAAAAAAGAAAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAA 1185

QY 140 CTAGTGACTTAGTAAAGGATATTTTAGCAGCACAAACTTAACAACAGGTGCAGTTATTT 199
DB 1184 AAAAAAATAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1125

QY 200 TGAACAAAGATACAAAGTTACTTTCTATGATGCCAAATGAGAAAGATTTCTTCACTCCAA 259
DB 1124 ATAAAAATATATAATAAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAA 1065

QY 260 CTGAGATAAAAAGTTTATTTCAGAACAACTTTAACTACAGCTAATGGAATGAAGATT 319
DB 1064 ATAAATATAAATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1005

```

```

QY 320 ATGTAAAGACAACTTTTAAAAAATTTAGATGCGAGGAGATATGCTATTATAGATTAACTT 379
DB 1004 AAATAAAA-----TATATAAATAAAAAAATATATAAATAAAGATAAAAAAGAAATAA 951

QY 380 ATAATAATGCTAAAACTGTTGAAATTTAAAGTAGTAGCAGCTAGTCAAAAAACAGTAGTTG 439
DB 950 TAAAAAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 891

QY 440 TATCTAGTAGTGCGAAAAATAGTCACAAAGATATATAGCTGAAAAATATATGTGTTTGAAGCA 499
DB 890 AAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 831

QY 500 AAGACTTAGAAAAATGCCTAAAAAATTAATAATGCTCAGATTTCTAGTAAAACTGATAGTT 559
DB 830 ATCAAGGGTAAAAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 771

QY 560 ACTATCAAGTAGTCTTTTATCCAAAGGAAGAGAGATTAACAAGGTTTCTCAACTTTATAGAG 619
DB 770 AAAAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 711

QY 620 CTACAAATTAATAATGAAGGAAGCTGCATATGTAATATACACAGTAAATATTAATCTTAAAT 679
DB 710 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 654

QY 680 CTACTAGTAAGAGTAAATTTAAAGACTGCGAGTAGAAGAGTTACAAAAATTTGAATGCTAGTT 739
DB 653 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 594

QY 740 ATTCTAATACTACAACCTTTAGCTGCTGATGACAGAGATACAAACAGCTATAGAGATAAGTA 799
DB 593 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 534

QY 800 AAGAATATTTACAATAATGATGCGGAGAAATCAGATCATTTCAAGCTGATGTTAAAGAGAAATG 859
DB 533 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 474

QY 860 TTAATAATGTTGTTATTTAGTAGTGCAATGCACTAGTAGTAGTAGTAGTGGGGTCTCTT 919
DB 473 AAAAAAATAAAGGGGAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 414

QY 920 TAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTTCAAAAGATAAATTAGATTTCGTCAG 979
DB 413 AAAAAAATAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 354

QY 980 TAAAAATCTGAATAAAGAGAGTTTTTAGACTTTAAAAAATTTCAACAGAGTAACAGGAAAAA 1039
DB 353 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 294

QY 1040 CAGTTTATATAGCTGGTGAGTTAATAGTGATCTTAAGAAGAGTTGTAACAGAAATTAGAAT 1099
DB 293 AACAATACAAAAAACAACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 234

QY 1100 CAATGGGATTAAGAGTTGAAAGATTTCTCAGGATGATAGATATGAACATCTCTTTAAAAA 1159
DB 233 ACTTTACACTTTTATCTACTTAAACAAATCATATATTTGTGAAATTTGAAACGGATAACAATT 174

QY 1160 TAGCAGGTGCAAAATAG 1174
DB 173 TCACACAGGAACAG 159

RESULT 12
CL118721
LOCUS CL118721
DEFINITION ISBI-7238_T7.1 ISBI Xenopus tropicalis genomic clone ISBI-7238,
genomic survey sequence.
ACCESSION CL118721
VERSION CL118721.1 GI:40612356
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1608)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 252
High quality sequence stop: 345.
Features
Location/Qualifiers
1. 1608
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone_lib="ISB1-72J8"
/notes="Vector: pBelBAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 5.6%; Score 101.8; DB 9; Length 1608;
Best Local Similarity 45.6%; Pred. No. 2.8e-10;
Matches 483; Conservative 0; Mismatches 562; Indels 14; Gaps 3;
QY 113 ATAGTATACAGTATCAATACATAAGCTAGTCACTTAGTAAAGGATATTTTACGACGAC 172
DB 541 AAAAAAATAAATAAATAAATCCCAATATATTTTAAAAAATAAATAAATAAATAA 600
QY 173 AAAAATTAAACACAGGTCAGTATTTTGAACAAAGATACAAAAGTTTCTTCTATGAT 232
DB 601 TATATAAAGAAATATAAATAATTATATTAAGATTATATAAATAAATAAATAAATA 660
QY 233 CAATGAGAAAGATCTTCAACTCCAACTCGAGATAAAGTTTATTCAGACAACTT 292
DB 661 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 720
QY 293 TAACACAGCTAATGGAATGAGATTATGTAAAGACAACTTTAAATAAATTTAGATGAC 352
DB 721 TATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 780
QY 353 GAGAAATATCTATTATAGATTAACTTTATATATATGCTAAACTGTTGAAATTAAGTAG 412
DB 781 AATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
QY 413 TAGCAGCTAGTGAATAAACAAGTAGTGTCTAGTATGCGAATAAATAGTGCATAAGATA 472
DB 841 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 900
QY 473 TAGCTGAATAATATGCTTTTGAACACAAAGCTTAGAATAATGCACTAAATACTATAATG 532
DB 901 TAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
QY 533 CCTCAGATTTTCAGTAAACTGATAGTTACTATCAAGTAGTTCTTTATCCAAAGGAAAGA 592
DB 961 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1017
QY 593 GATTACAGGTTCTTCAACTTATAGAGCTACAATAATTATATGAAGGAACTGCATATGCTA 652
DB 1018 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1077
QY 653 ATACACAGTATATTAATTAACCTTAATACTAGTATAGAGTAATTTT-----AAAGA 703
DB 1078 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1137
QY 704 CTGCAGTAGAAGAGTTACAAAAATTTGAATGCTAGTATTTCTAATATACAACTTTAGCTG 763

DB 1138 AAATATTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1197
QY 764 GTGATGACAGATACAAACAGCTATAGAGATAAGTAAAGAAATATTACAAATATGATGGCG 823
DB 1198 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1257
QY 824 AGAAATCAGATCATTCAGCTGATGTTAAAGAGAAATGTTAA--AAATGTTGTTATTAGTAGG 881
DB 1258 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1317
QY 882 TCCAATGCACCTAGTAGTGGATTAGTTGCGGCTCTCTTTAGCAGCAGAAAAAGAGCTCTCC 941
DB 1318 AAATAATATATTAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1377
QY 942 ACTATTATTAACTTCAAAAGATAAATAATTAGATTTCGTCAGTAAATCTGAAATAAAGAGAGT 1001
DB 1378 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1437
QY 1002 TTTGAGCTTAAAACTTCAACAGAGTAAACAGAAAAACAGTTTATATAGCTGGTGGAGT 1061
DB 1438 GAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1497
QY 1062 TAATAGTGTATCTAAAGAGTTGTAAACAGAAATTTAGAAATCAATGGGATTTAAAAAGTTCGAAAG 1121
DB 1498 AAATAAATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1557
QY 1122 ATTCTCAGCTGATGATAGATATGAAACTTCTTTAAAAAT 1160
DB 1558 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1596
CL078613 1594 bp DNA linear GSS 31-DEC-2003
CH216-151M17_Sp5.1 CH216 Xenopus tropicalis genomic clone
CH216-151M17, genomic survey sequence.
CL078613
CL078613.1 GI:40534526
GSS.
Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert length: 175000 Std Error: 0.00
Seq primer: Sp5 atctgcgcttcgaccc
Class: BAC ends
High quality sequence start: 557
High quality sequence stop: 634.
Features
Location/Qualifiers
1. 1594
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-151M17"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/notes="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"
ORIGIN
Query Match 5.6%; Score 101.6; DB 9; Length 1594;

```
Best Local Similarity 42.5%; Pred. No. 3.1e-10;
Matches 470; Conservative 0; Mismatches 634; Indels 2; Gaps 1;

QY 87 AATATCACTACAGATGGTACAAATGATAGTATACAGATATCAAAATACATAAGCTAGTGA 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 409

QY 147 CTTAGTAAAGGATATTTTAGCAGCACAAACTTAAACAAGGTGCAGTTATTTTGAACAA 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 NNAANANANANANANANANANANANANANANANANANANANANANANANANANAN 469

QY 207 AGATACAAAGGTCTCTATGATGCAATGCAAGAGATCTTCAACTCCAAGTGGAGA 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 529

QY 267 TAAAAAGTTTATTCAGAACAACTTTAACTACAGCTAATGAAATGAAGATTATGTAAA 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 589

QY 327 GACAACTTTAAATAATTTAGTCAGGAGAATATGCTATATAGATTTAACTTTAATAA 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 649

QY 387 TGTAAACTGTTGAAATTAAGTACTAGCAGCTAGTGAAGAGACAGTAGTTGTCTAG 446
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 709

QY 447 TGATGCGAAATATAGTGCAGAAAGATATAGCTGAAATATATGTTTTGAAGACAAAGCTT 506
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 710 AANNNANANANANANANANANANANANANANANANANANANANANANANANAN 769

QY 507 AGAAATGCACTAAACCTATAAATGCTCAGATTTTCAGTAAACCTGATAGTTACTATCA 566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 770 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 829

QY 567 AGTAGTCTTTATCCAAAGGAAGAGATTACAAAGTTTCTCAACTTATAGAGCTACAAA 626
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 830 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 889

QY 627 TTATAATGAAGGAATCGCATATGTTAATACACAGTAATATTTAACTC--TAAAACTACT 684
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 890 AAAAAAAAAANANATTAANANANANANANANANANANANANANATTAANANATTAANAN 949

QY 685 AGTAAGAGTAATTTAAAGACTGCGATGAGAGAGTTACAAAATTTGAATGCTAGTTATCT 744
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 950 AAAAAAAAAANATTAANANANANANANANANANANANANANATTAANANATTAANAN 1009

QY 745 AATACTACAACTTTAGCTGCTGATGACAGATACAAACAGCTATACAGATAAGTAAAGAA 804
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1010 AAAAAAAAAAAAAAAAAANANANANANANANANANANANANANANATTAANANATTAAN 1069

QY 805 TATTACAATTAATGATGCGGAGAAATCAGATCATTTCAAGCTGATGTTTAAAGAGAAATGTTAA 864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1070 AATATAAAAAAAAAAAAAAAAAAAAAAAAAANANATTAANANANANANANANANANAN 1129

QY 865 AATGTTGTTATAGTAGTGAATGCACTAGTAGAGATGATTTAGTTCGGCTCCTTTAGCA 924
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1130 AAAAAAAAAAAAAAAAAATTAANANANANANATTAANANATTAANANANANANANAN 1189

QY 925 GCAGAAAAAGATGCTCCACTATTATTAACTTCAAAAGATAAATTTAGATTTCGTCAGTAAA 984
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1190 AAAAAATTAANANANANANANANANANANANANATTAANANANANANANATTAANAN 1249

QY 985 TCTGAAATTAAGAGAGTTTTAGACTTTAAAACTTCAACAGAGTAACAGAGAAAAACAGTT 1044
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1250 AATATAAAAAAAAAAAAAAAAAATTAANANATTAANANANANANANANANANANATTA 1309

QY 1045 TATATAGCTGGTAGTTAATAGTGATCTAAAGAGAGTTGTTAACAAGATTTAGAAATCAATG 1104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1310 AAAAAAAAAATTAANANATTAANANANANANATTAANANANANANANANANANAN 1369

QY 1105 GGATTAAGAGTTCTCAAGGATTCAGGTGATGATAGATATGAAACTCTTTTAAANATAGCA 1164
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1370 AAAAAAAAAATAAANAAAAAAAAAAAAAAAAATAAAAAAAAAATAAAAAAAAAAAAAA 1429

QY 1165 GGTGAATAGGCTTAGATAATGATAA 1190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1430 AATTAATAAAAAAAAAATAAATAA 1455

RESULT 14
CL0771122 1162 bp DNA linear GSS 31-DEC-2003
LOCUS CH216-143B5 Sp6.1 CH216 Xenopus tropicalis genomic clone
DEFINITION CH216-143B5_ genomic survey sequence.
ACCESSION CL0771122
VERSION CL0771122.1 GI:40533035
SOURCE GSS.
ORGANISM Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1162)
AUTHORS Krenitski,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 757
High quality sequence stop: 808.
Location/Qualifiers
1. .1162
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strains="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-143B5"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 5.5%; Score 100.4; DB 9; Length 1162;
Best Local Similarity 42.7%; Pred. No. 5.3e-10;
Matches 471; Conservative 0; Mismatches 630; Indels 3; Gaps 1;

QY 87 AATATCACTACAGATGGTACAAATGATAGTATACAGTATCAAAATACATAAGCTAGTGA 146
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 AATCTCCAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 97

QY 147 CTTAGTAAAGGATATTTTAGCAGCACAAACTTAAACAAGGTGCAGTTATTTTGAACAA 206
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 AAAAAAAAAAATCTTAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 157

QY 207 AGATACAAAGTTACTTTCTATGATGCAAAATCAGAAAGATTCTTCAACTCCAAGTGGAGA 266
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 AAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 217

QY 267 TAAAAAGTTTATTCAGAACAACTTTTAACTACAGCTATGGAATGAAGATTATGTAAA 326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 AAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 277

QY 327 GACAACTTTAAATAATTTAGATGCGAGAGAAATATGCTATATAGATTTAACTTTAATAA 386
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 AAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 337

QY 387 TGTAAACTGTTGAAATTTAAAGTAGTAGCACCTAGTGAAACACAGTAGTTGTATCTAG 446
```

```

338 AAAAAAAAAAAAAAAAAAATAATTTGAAAAAAGAAAAAAAAAAAAAAAAAATTTAAAAAAGAAAAA 397
447 TGATGCCAAAAATAGTCCAAAGATATAGCTGAAAAATATGTGTTTGAAGACAAAGACTT 506
398 AAAAAAAAAAAAAAAAAAATAATTTAAAAAAGAAAAAAAAAAAAAAAAAATTTAAAAAAGAAAAA 457
507 AGAAATGCACTAAAAAATAAATGCTCAGATTTTCAGTAAAGCTGATAGTTACTATCA 566
458 AAAAAAAAAAAAAAAAAAATAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 517
567 AGTAGTTCTTTATCCAAAGGAAAGAGATTACAAAGGTTTCTCACTTTATAGAGCTACAAA 626
518 AAAAAAAAAAAAAAAAAAATAAAGAAAAAATAAAGAAAAAATAAAGAAAAAATAAAGAAAAA 577
627 TTATATGAAGGAAGCTCATATGTTATATACACAGTAAATTTAACTTAAATCTTACTAG 686
578 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 637
687 TAGAGTAATTTAAAGACTGCAGTAGAAGAGTTACAAAAATTTGAATGCTAGTTATCTAA 746
638 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 697
747 TACTACAACTTTAGCTGGTGATGACAGATACAAACAGCTATAGAGATAAGTAAAGATA 806
698 AAAAAAAAAAATAAAGAAAAAATAAAGAAAAAATAAAGAAAAAATAAAGAAAAAATAAAG 757
807 TTCAATATATGATGGCAGAAATCAGATCAATCAGCTGATGTTTAAAGAGATGTTAAAAA 866
758 AAAAAAAAAAATAAAGAAAAAATAAAGAAAAAATAAAGAAAAAATAAAGAAAAAATAAAG 817
867 TGTTGTTATTAGTAGTCAATGCACTAGTAGATGAGTATAGTTAGTCGCTCTCTTTAGCAGC 926
818 ATATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 877
927 AGAAAAAGATGCTCCACTATTATTAACTTCAAAAGATAAATAAATAAATAAATAAATAAATA 986
878 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 937
987 TGAATTAAGAGAGATTTTAGCTTAAATACTTCAACAGAGTACAGGAAAAACAGTTTA 1046
938 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 997
1047 TATAGCTGGTGGATTATAGTGTATCTTAAAGAGATTTGTAACAGATTTAGATCAATGGG 1106
998 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1054
1107 ATAAAAAGTTCGAAAGATTTCTCAGGTGATGATAGATATGAAACTTCTTTAAAAATAGCAGG 1166
1055 AATAATATGATAAATAAATAAAGAAAGATTAATAAATAAATAAATAAATAAATAAATAAATA 1114
1167 TGAATATAGCTTAGAATATGATAA 1190
1115 TTAGAAAAATAAATAAATAAATAA 1138

RESULT 15
AG386893/c
LOCUS
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-201E07.TJ, genomic survey
sequence.
ACCESSION AG386893.1 GI:47998098
VERSION AG386893.1
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
UNPUBLISHED
REFERENCE 2 (bases 1 to 1519)

```

```

AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Taukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
LIBRARY Sequencing : TJ
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source
1..1519
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-201E07.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match 5.5%; Score 100.4; DB 9; Length 1519;
Best Local Similarity 40.7%; Pred. No. 5.4e-10;
Matches 451; Conservative 0; Mismatches 657; Indels 1; Gaps 1;
Qy 83 ATGTAATATACATCAAGATGGTACAAATGATAGTATACAGTATCAATCAATCAAGCTA 142
Db 1418 ANANANANANANANANANANANANANANANANANANANANANANANANANANANANAN 1359
Qy 143 GTGACTTAGTAAGGATATTTTACGACACAAACTTAAACAAAGGTGCGAGTTATTTGA 202
Db 1358 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANANAN 1299
Qy 203 ACAAAGATACAAAGATTTACTTCTATCATGCAAAATGAGAAAGATTTCTCAACTCCAAG 262
Db 1298 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 1239
Qy 263 GAGATAAAAAAGTTTATTCAGAAACAACTTTTAACCTACAGCTAATGGAATGGAAGATT 322
Db 1238 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 1179
Qy 323 TAAAGACAACTTTAAAAAATTTAGATCGAGGAGATATGCTATTATAGATTTAATCTATA 382
Db 1178 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 1119
Qy 383 ATATGCTAAAGCTGTTGAAATTAAGTAGTAGCTAGTAGAGAAAAACAGTAGTTGTAT 442
Db 1118 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 1059
Qy 443 CTAGTGATGCGAAAAATAGTGCAAAAGATATAGCTGAAAAATAT-GTGTTTGAAGACAAA 501
Db 1058 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 999
Qy 502 GACTTAGAAAAATGCACTTAAAAACTATAATGCCTCAGATTTTCAGTAAAGCTGATGTAC 561
Db 998 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 939
Qy 562 TATCAAGTAGTTCTTTTATCCAAAGGAAAGAGATTTACAAGGTTTCTTCACTTATAGACT 621
Db 938 AAAAAANANANANANANANANANANANANANANANANANANANANANANANANAN 879
Qy 622 ACAATTTATATGAAGGAAGTGCATATGGTAAATACACCATGTAATTTAACTCTAAATCT 681

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM nucleic - nucleic search, using sw model

Run on: October 6, 2005, 06:06:36 ; Search time 3690 Seconds
(without alignments)
3454.108 Million cell updates/sec

Title: US-10-068-870A-4
Perfect score: 1830
Sequence: 1 atgaaaaaagaatttagc.....tgaagattattagatatg 1830

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 8443130 seqs, 3482420727 residues

Total number of hits satisfying chosen parameters: 16886260

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1830	100.0	1830	14	US-10-068-870-2
2	542	29.6	2271	14	US-10-068-870-4
3	535.6	29.3	2271	18	US-10-239-610-6
4	503.4	27.5	2217	14	US-10-068-870-6
5	483.2	26.4	2145	18	US-10-239-610-4
6	480	26.2	2145	14	US-10-068-870-7
7	471.8	25.8	2157	14	US-10-068-870-1
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 6, Appli
					Sequence 6, Appli
					Sequence 4, Appli
					Sequence 7, Appli
					Sequence 1, Appli

8	471.8	25.8	2158	14	US-10-068-870-3	Sequence 3, Appli
9	471.8	25.8	2158	14	US-10-068-870-8	Sequence 8, Appli
10	471.8	25.8	2160	18	US-10-239-610-5	Sequence 5, Appli
11	470.2	25.7	2158	14	US-10-068-870-5	Sequence 5, Appli
12	145.2	7.9	1893	17	US-10-282-122A-16859	Sequence 16859, A
13	104	5.7	1503	17	US-10-282-122A-16782	Sequence 16782, A
14	99.6	5.4	1830	17	US-10-282-122A-16845	Sequence 16845, A
15	95.6	5.2	5361	9	US-09-742-096-2	Sequence 2, Appli
16	95.6	5.2	6152	9	US-09-742-096-1	Sequence 1, Appli
17	94.8	5.2	5528	18	US-10-415-253-1	Sequence 1, Appli
18	92	5.0	4997	17	US-10-282-122A-35506	Sequence 35506, A
19	88.4	4.8	5253	17	US-10-282-122A-17449	Sequence 17449, A
20	87.6	4.8	7442	18	US-10-221-714A-409	Sequence 409, App
21	86.2	4.7	7446	17	US-10-282-122A-7578	Sequence 7578, Ap
22	85.2	4.7	7434	9	US-09-815-242-4761	Sequence 4761, Ap
23	85.2	4.7	7437	9	US-10-470-048B-201	Sequence 201, App
24	85.2	4.7	8155	8	US-09-815-242-8869	Sequence 8869, Ap
25	85.2	4.7	8155	8	US-08-781-986A-63	Sequence 63, Appl
26	85.2	4.7	8155	18	US-10-329-624-63	Sequence 63, Appl
27	84.8	4.6	1300	19	US-10-668-749A-1	Sequence 1, Appli
28	84.2	4.6	4911	21	US-10-470-048B-427	Sequence 427, App
29	82.2	4.5	2370	11	US-09-754-947-2	Sequence 2, Appli
30	81.8	4.5	6292	18	US-10-221-714A-461	Sequence 461, App
31	81.4	4.4	1243	20	US-10-425-115-172717	Sequence 172717,
32	80.6	4.4	2409	17	US-10-282-122A-17248	Sequence 17248, A
33	79	4.3	1204	19	US-10-437-963-77858	Sequence 77858, A
34	76.4	4.2	921	20	US-10-425-115-38710	Sequence 38710, A
35	74.8	4.1	778	20	US-10-363-345A-2179	Sequence 2179, Ap
36	74.8	4.1	778	20	US-10-363-483A-2179	Sequence 2180, Ap
37	74.8	4.1	778	21	US-10-363-483A-2179	Sequence 2180, Ap
38	74.8	4.1	1039	20	US-10-363-483A-2180	Sequence 2180, Ap
39	74.8	4.1	1039	20	US-10-425-115-58802	Sequence 58802, A
40	74.2	4.1	50000	21	US-10-706-635-25	Sequence 25, Appl
41	74	4.0	1062	20	US-10-425-115-120013	Sequence 120013,
42	73.8	4.0	15548	15	US-10-311-455-2128	Sequence 2128, Ap
43	73.8	4.0	3673778	16	US-10-312-841-1	Sequence 1, Appli
44	73.6	4.0	1845	17	US-10-282-122A-16828	Sequence 16828, A
45	73.2	4.0	1830	21	US-10-706-635-67	Sequence 67, Appl

ALIGNMENTS:

RESULT 1

US-10-068-870-2
; Sequence 2, Application US/10068870
; Publication No. US20030054009A1
; GENERAL INFORMATION:
; APPLICANT: The Provost, Fellows & Scholars of the College of the Holy and Undivided
; APPLICANT: Trinity of Queen Elizabeth, near Dublin
; TITLE OF INVENTION: C.difficile vaccine
; FILE REFERENCE: TRI002/C/WO
; CURRENT APPLICATION NUMBER: US/10/068,870
; CURRENT FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Clostridium difficile
; US-10-068-870-2

Query Match 100.0%; Score 1830; DB 14; Length 1830;
Best Local Similarity 100.0%; Pred. No. 1.3e-277;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAAAAAAGAAATTTAGCAATGGCTATGCGAGCTGTACTGTAGTAGGTTCTGCTGCT	60
Db	1	ATGAAAAAAGAAATTTAGCAATGGCTATGCGAGCTGTACTGTAGTAGGTTCTGCTGCT	60
Qy	61	CCAGTTTTTGCAGCAGCTTCAGATGTAATATCATCAAGATGCTACAAATGATAAGTAT	120
Db	61	CCAGTTTTTGCAGCAGCTTCAGATGTAATATCATCAAGATGCTACAAATGATAAGTAT	120

QY 121 ACAGTATCAAACTAAAGCTAGTGAAGTAAAGGATATTTTACGACGACAAAACCTTA 180
 Db 121 ACAGTATCAAACTAAAGCTAGTGAAGTAAAGGATATTTTACGACGACAAAACCTTA 180
 QY 181 ACAACAGGTGCAGTATTTTGAACAAAGATACAAAAGTTACTTTCTATGATGCAAAATGAG 240
 Db 181 ACAACAGGTGCAGTATTTTGAACAAAGATACAAAAGTTACTTTCTATGATGCAAAATGAG 240
 QY 241 AAGATATCTTCAACTCAACTCGAGATAAAAAAGTTTATTCAGAACAAAACCTTTAACTACA 300
 Db 241 AAGATATCTTCAACTCAACTCGAGATAAAAAAGTTTATTCAGAACAAAACCTTTAACTACA 300
 QY 301 GCTAATGGAAATCAAGATTTATGTAAGACAACTTTTAAAGATTTAGATCGAGAGATAT 360
 Db 301 GCTAATGGAAATCAAGATTTATGTAAGACAACTTTTAAAGATTTAGATCGAGAGATAT 360
 QY 361 GCTATTATAGATTTAACTTTATATAATGCTTAAACCTGTTGAAATTAAGTATGAGCAGCT 420
 Db 361 GCTATTATAGATTTAACTTTATATAATGCTTAAACCTGTTGAAATTAAGTATGAGCAGCT 420
 QY 421 AGTGAAAAACAGTATGTTGATCTAGTATGCGAAAAATAGTGCAAAAGATATAGCTGAA 480
 Db 421 AGTGAAAAACAGTATGTTGATCTAGTATGCGAAAAATAGTGCAAAAGATATAGCTGAA 480
 QY 481 AATATGCTGTTGAACACAAAGACTTAGAAAATGCACTTAAACACTTAAATGCTCAGAT 540
 Db 481 AATATGCTGTTGAACACAAAGACTTAGAAAATGCACTTAAACACTTAAATGCTCAGAT 540
 QY 541 TTCAAGTAAAACTGATAGTTACTATCAAGTAGTTCTTTTATCCAAAAGGAAAGATATACAA 600
 Db 541 TTCAAGTAAAACTGATAGTTACTATCAAGTAGTTCTTTTATCCAAAAGGAAAGATATACAA 600
 QY 601 GGTTTCTCAACTTATAGAGCTACAAATATATGAAGGAAGTGCATATGGTAATACACCA 660
 Db 601 GGTTTCTCAACTTATAGAGCTACAAATATATGAAGGAAGTGCATATGGTAATACACCA 660
 QY 661 GTATATTAACCTTAAATCTACTAGTATAGAGTAAATTTAAAGCTCAGTAGAAGGTTA 720
 Db 661 GTATATTAACCTTAAATCTACTAGTATAGAGTAAATTTAAAGCTCAGTAGAAGGTTA 720
 QY 721 CAAAAATGGAATGCTAGTTATTTCTAATACCTACAACTTTAGCTGGTGCATGACAGAAATACAA 780
 Db 721 CAAAAATGGAATGCTAGTTATTTCTAATACCTACAACTTTAGCTGGTGCATGACAGAAATACAA 780
 QY 781 ACAGCTATAGAGATAAGTAAAGAAATATCAATATGATGGCGAGAAATCAGATCATTC 840
 Db 781 ACAGCTATAGAGATAAGTAAAGAAATATCAATATGATGGCGAGAAATCAGATCATTC 840
 QY 841 GCTGATGTTAAAGAGAAATGTTAAAGATTTGTTATTTAGTATAGTGGCAATGCACTAGTAGAT 900
 Db 841 GCTGATGTTAAAGAGAAATGTTAAAGATTTGTTATTTAGTATAGTGGCAATGCACTAGTAGAT 900
 QY 901 GGATTTAGTTGCGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTTAACTTCAAAA 960
 Db 901 GGATTTAGTTGCGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTTAACTTCAAAA 960
 QY 961 GATAAATTTAGATTCGTCAGTAAATCTGAAATTAAGAGAGTTTATAGACTTAAAGAACTTCA 1020
 Db 961 GATAAATTTAGATTCGTCAGTAAATCTGAAATTAAGAGAGTTTATAGACTTAAAGAACTTCA 1020
 QY 1021 ACAGAGTAAACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAAATAGTGTATCTAAGAA 1080
 Db 1021 ACAGAGTAAACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAAATAGTGTATCTAAGAA 1080
 QY 1081 GTTGTAACAGAAATTTAGAACTCAATGGGATTTAAAGATTTGAAAGATTTCTCAGGTGATGATAGA 1140
 Db 1081 GTTGTAACAGAAATTTAGAACTCAATGGGATTTAAAGATTTGAAAGATTTCTCAGGTGATGATAGA 1140
 QY 1141 TATGAAACCTTTTAAAAATAGCAGGTGAAATAGGCTTTAGATAATGATAAGGCTTATGTA 1200
 Db 1141 TATGAAACCTTTTAAAAATAGCAGGTGAAATAGGCTTTAGATAATGATAAGGCTTATGTA 1200

QY 1201 GTTGGTGGAAACAGGATTTAGCAGATGCCATGACTATAGCTTCAAGTTGCTTCTACTAAATTA 1260
 Db 1201 GTTGGTGGAAACAGGATTTAGCAGATGCCATGACTATAGCTTCAAGTTGCTTCTACTAAATTA 1260
 QY 1261 GATGGTAAATGGTGTGTTAGATAGAAACAAATGACATGCTTACTCCAAATAGTTCTTGTGTAGAT 1320
 Db 1261 GATGGTAAATGGTGTGTTAGATAGAAACAAATGACATGCTTACTCCAAATAGTTCTTGTGTAGAT 1320
 QY 1321 GGAAAGCTGATGATAAAATATCTGATGACTTAGATAGTTTCTTAGGAAGCGCTGATGTAGAT 1380
 Db 1321 GGAAAGCTGATAAAATATCTGATGACTTAGATAGTTTCTTAGGAAGCGCTGATGTAGAT 1380
 QY 1381 ATAAATAGTGGATTTGCAAGTCTATCTGAAAGATGGAAGAGCTATATCAGATGCTACT 1440
 Db 1381 ATAAATAGTGGATTTGCAAGTCTATCTGAAAGATGGAAGAGCTATATCAGATGCTACT 1440
 QY 1441 GGTAAAGGCGTTTACAAGAGTTTAAAGCGCACATAGACAAGACATACTCTGAAGTTATA 1500
 Db 1441 GGTAAAGGCGTTTACAAGAGTTTAAAGCGCACATAGACAAGACATACTCTGAAGTTATA 1500
 QY 1501 AAAACATATATGCTTAATGATATCTGAAATAGCTGCAAGTGTCTTTTAGATAAAGATTCA 1560
 Db 1501 AAAACATATATGCTTAATGATATCTGAAATAGCTGCAAGTGTCTTTTAGATAAAGATTCA 1560
 QY 1561 GGTGCTTCAAGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 1561 GGTGCTTCAAGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1621 AAGAAGATCAATTAAGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 1621 AAGAAGATCAATTAAGTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1681 CAGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Db 1681 CAGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 QY 1741 GTAGAGAAAAATATTTCTAAAGATTTAAACAGTTTAAAGATTTAAAGATTTAAAGATTTAAAGAT 1800
 Db 1741 GTAGAGAAAAATATTTCTAAAGATTTAAACAGTTTAAAGATTTAAAGATTTAAAGATTTAAAGAT 1800
 QY 1801 ATAAACAAAATGAAAGATTTTATTAGATATG 1830
 Db 1801 ATAAACAAAATGAAAGATTTTATTAGATATG 1830

RESULT 2

US-10-068-870-4
 ; Sequence 4, Application US/10068870
 ; Publication No. US20030054009A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Provost, Fellows & Scholars of the College of the Holy and Undivided
 ; APPLICANT: Trinity of Queen Elizabeth, near Dublin
 ; TITLE OF INVENTION: C.difficile vaccine
 ; FILE REFERENCE: TRI002/C/NO
 ; CURRENT APPLICATION NUMBER: US/10/068,870
 ; CURRENT FILING DATE: 2002-02-11
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 2271
 ; TYPE: DNA
 ; ORGANISM: Clostridium difficile
 US-10-068-870-4

Query Match 29.6%; Score 542; DB 14; Length 2271;
 Best Local Similarity 59.4%; Pred. No. 1.6e-75;
 Matches 1117; Conservative 0; Mismatches 670; Indels 93; Gaps 8;
 QY 32 CAGCTGTTACTGCTAGTAGGTTCTGCTCCAGTTTTCAGCAGCAGCTTCAGATGTAATAT 91
 Db 401 CAGCAGATGCTATATAATGCTGGAAATCTTCAGCAGATGGTGTGTTTACAAATACTGGAG 460
 QY 92 CACTACAAGATGGTCAAAATGATATACAGTATCAAACTACTAAAGCTAGTACTTAG 151

Db 461 CTGCTAGTGGTCTCTAGAGACAAATTCAGCAGGAACAAAATCTGCAATGTCAGCTATTT 520
 QY 152 TAAAGGATATTTAGCAGACAAAATTAACACAGGTGCAGTTATTTTGAACAAGATA 211
 Db 521 TTGACACAGCATATACAGATTCATCTGAACTGCGGTTAAGATTTACTATAAAGCAGATA 580
 QY 212 CAAAGGTACT---TTCTATGATGCAAAATGAGAAAGATTTCTTCAACTCCAACTGGAGATA 268
 Db 581 TGAATGATTAATAATTTGGTAAAGCAGGTGAGACAACTTATTCACCTGGGCTTACATTG 640
 QY 269 AAAAGCTTTATTCAGAACAACTTTAACTACAGCTAATGGAATGAAGATTAATGTHAAGA 328
 Db 641 AAGATGGCTCTACAGAAAAAATTTGTTAAATTTAGGGACAGTGATATATATAGATATACTA 700
 QY 329 CAACTTTAAA---AAATTTAGTCAGGAGATATGCTATATATAGATTTAACTTAATA 385
 Db 701 AAGCTCTTAACTTACTTGTGTTCTGGAAGTAAAGCAACTGTTAAGTTTGTGGAANAAA 760
 QY 386 ATGCTAAAACGTGTG-----AAATTTAAGTGTAGCAGCTAGTGAAA 427
 Db 761 CACCAAGTGCAGTGTCAACAGTAATAACAAGCTTTAGAATAATAATGCTTAAAGAG 820
 QY 428 AAACAGTGTGTATCTAGTGTGCGAAAAATAGTCCAAAAGATATAGCTGAAAAATATG 487
 Db 821 AAACAATAGATATTTAGCGCTAGTTCTAGTAAAAACAGCACAAGATTTAGCTTAAAAATATG 880
 QY 488 TGTGTAAGCAAGAAGCTTTAGAAATGCACTAAAACTATAAATGCTCAGATTTTCAGTA 547
 Db 881 TATTTAATAAACTGATTTAAATACTCTTTTATAAAGTATTAAATGGAGATGAAGCAGATA 940
 QY 548 AAACATGATGTTACTATCAAGTAGTTCTTTATCCAAAAGCAAGAGATTACAAGTTTCT 607
 Db 941 CTAAATGATTAATAGAAAGATTTAGTGGAAAAATATCAAGTAGTTCTTTATCCAGAGGAA 1000
 QY 608 CAACCTTATAGAGCTACAAATTTAT---AATGAAGGAAGCTGCATATGTTAATPACACAGTAA 664
 Db 1001 AAGAGTTACAACTAAGAGTGTGCAAAAGCTTCAATTTGCTGATGAAAAATCCACCAGTTA 1060
 QY 665 TATTAACCTAAATCTACTAGTAAGAGTAATTTAAGACATGCAAGTGAAGAGATTACAAA 724
 Db 1061 AATTAACCTTAACTCAGATAAAGAAAGAACTTAAAGATTAATGTTGATGATTTAAGAA 1120
 QY 725 AATGATGCTAGTTATTTCTAATCTACAACTTTAGCTGTGATGACAGAAATCAACACAG 784
 Db 1121 CATATAAATAGGATTTCAATGCTATAGAAGTACGAGAGAGATAGAAATAGAACTG 1180
 QY 785 CTATAGAGATAAGTAAAGAAATTTACAAATTAATGATGGCGAGAAATCAGATCATTCAGCTG 844
 Db 1181 CAATAGCATTAAGTCAAAATAATTTAATCTCTGATGATGAAA-----TGCTA 1228
 QY 845 ATGTTAAGAGAAATGTTAAAAATGTTGATTTAGTAGTGCAATGCACTAGTAGATGAT 904
 Db 1229 TATTTAGAGATTCAGTTTGATAATGTAGTATTTGTTGGAGGAAATGCAATAGTTGATGGAC 1288
 QY 905 TAGTTGCGGCTCTTTAGCAGCAAAAAGATGCTCCACTATTATTAATTAATCTCAAAAGATA 964
 Db 1289 TTGTAGCTTCTCTTGTAGCTTCTGAAAAGAAAGCTCTCTTTATTAATTAATCTCAAAAGATA 1348
 QY 965 AATTAGATTCGTAGTAAATCTGAAAATAAAGAGAGTTTGTAGCTTAAAACTTCAACAG 1024
 Db 1349 AATTAGATTCNAGCTAAAGCTGAATTAAGAGAGTTATGAATATAAAGAGTACAACAG 1408
 QY 1025 AAGTA---ACAGGAAAAAGCTTTATATAGCTGTGGAGTTAATAGTGTATCTTAAGAAG 1081
 Db 1409 GTATAAATACTTCAAGAAAAATTTATTTAGCTGTGGAGTTAATTTCTATATCTTAAGAAG 1468
 QY 1082 TTGTAACAGATTTAGATCAATGAGTTTAAAGTTGAAGATTTCTCAGGTGATCATAGAT 1141
 Db 1469 TAGAAAAATGAATTAAGATATGGACCTTAAAGTTTACAAGATTAAGCAGGAGATGATAGAT 1528
 QY 1142 ATGAACTCTTTTAAAAATAGCAGTGAAATAGGCTTAGATAAATGATAAGGCTTTATGTAG 1201

Db 1529 ATGAAACTTCTCTAAAAATAGCTGATGAAGTAGGTTCTTGATTAATGATAAAGCATTTGTAG 1588
 QY 1202 TTGTTGGGAACAGGATTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTCTCTAATAATTAG 1261
 Db 1589 TTGGAGGAACAGGATTAGCAGATGCCATGAGTATAGCTCCAGTTGCTCATCTCAATTAAGAA 1648
 QY 1262 ATGGTAATGCTGTTGTAGATAGAAACAATGGACATGCTACTCCAACTAGTTGTTGTAGATG 1321
 Db 1649 ATGCTAATGTTGTAATTTAGCTGATGGTGCTACACCAATAGTAGTTGTAGATG 1708
 QY 1322 GAAAAGCTGATAAAATATCTGATGACTTAGATAGTTCTTTAGGAAGCGCTGATGTAGATA 1381
 Db 1709 GAAAAGCTTAAACTATTAATGATGATGTAAGAGATTTCTTAGATGATTCACAAGTTGATA 1768
 QY 1382 TAATAGTGGATTTGCAAGTGTATCTGAAAAGATGGAAGAAGCTATATCAGATGCTACTG 1441
 Db 1769 TAATAGTGGAGAAAACAGTGTATCTTAAAGATGTTGAAAATGCAATAGATGATGCTACAG 1828
 QY 1442 GTAAAAGCGCTTACAGAGTTTAAAGGCGAGATAGACAAGACACTACTCTGAAAGTTATAA 1501
 Db 1829 GTAAAATCTCCAGATAGATATAGTGGAGATGATAGACAAGCAACTAATGCAAAAAGTTATAA 1888
 QY 1502 AAACATATTTATGCTAATGATGCTGAAAATAGCTTAAAGCTGCAGTTTGTAGATAAAGATTG 1561
 Db 1889 AAGAATCTTCTTATTAATCAAGATACTTAATTAATGATAAATAAAGTAGTTAATTTCTTTG 1948
 QY 1562 GTGCTTCAAGTAGTAGTGCAGGAGTATTTAATTTCTATGTAGCTAAAGATGGAATCTACAA 1621
 Db 1949 TAGCTAAAGATGGTTCTTACTTAAAGAGATCAATTTAGTTGATGCTTTTAGCAGCAGCTCCAG 2008
 QY 1622 A-----AGAGATCAATTAGTTG 1639
 Db 2009 TTGCAGCAAACTTTGTTGTAACCTTTAATTTCTGATGTAAGCCAGTAGATAAAGATGGTA 2068
 QY 1640 ATGCATTAGCAGTAGGAGCTGTTGCTGGATATAA-----CTTGCTCCAGTTGTAT 1690
 Db 2069 AAGTATTAATGCTGTTCTGATTAATGATAAATAAATAGTATCTCCAGCACCTTAGTAT 2128
 QY 1691 TAGCTACTGATCTTTTATCTTCTGATCAATCGGTTGCTATAGCAAAAGTTGTAGGAGAAA 1750
 Db 2129 TAGCTACTGATCTTTTATCTTCTGATCAAAAGTGTATCTATAGTAAAGTTCTTGATAAAG 2188
 QY 1751 AATATTCTAAAGATTTAAACAAGTTGGTCAAGGATAGCTAATTCAGTTTATAAACAATA 1810
 Db 2189 ATAATGAGAAAACTTAGTTCAAGTTGGTAAAGGTATAGTACTTTCAGTTTATAAACAATA 2248
 QY 1811 TGAAGATTTTATAGATAG 1830
 Db 2249 TGAAGATTTTATAGATAG 2268

RESULT 3
 US-10-239-610-6
 ; Sequence 6, Application US/10239610
 ; Publication No. US20040039165A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Imperial College of Science Technology and Medicine
 ; TITLE OF INVENTION: CLOSTRIDIUM DIFFICILE POLYPEPTIDES AND USES THEREOF
 ; FILE REFERENCE: Y002 2 00089
 ; CURRENT APPLICATION NUMBER: US/10/239,610
 ; FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: PCT/GB01/01305
 ; PRIOR FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: GB 0007263.7
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 2271
 ; TYPE: DNA
 ; ORGANISM: Clostridium difficile
 US-10-239-610-6

Query Match		29.3%;	Score 535.6;	DB 18;	Length 2271;
Best Local Similarity		59.2%;	Pred. No. 1.6e-74;	Indels 93;	Gaps 8;
Matches 1113;		Conservative	0;	Mismatches 674;	
QY	32	CAGCTGTTACTGTTAGTGGTTCCTGCTCCAGTCTTTTTCGACGAGCTTCAGATGTAATAT	91		
DB	401	CAGCAGATGTAATAATGCTGGAAACATCTTCAGCAGATGGTGTGTTACAATACTGGAG	460		
QY	92	CACATAAAGTGTGTAACAATGATACTACAGTATCAAACTCAAAAGCTAGTGAATTAG	151		
DB	461	CTGCTAGTGGTCTACTGAGACAAATTCAGCAGAAACAACTTGAATGTGAGCTATTT	520		
QY	152	TAAAGGATTTTAGCAGCACAAACTTAAACAGGTGCAGTTATTTTGAACAAAGATA	211		
DB	521	TTGACACAGCATATACAGATTCATCTGAACTCGCGTTAAGATTACTATANAAGCAGATA	580		
QY	212	CAAAAGTTACT---TTCTATGATGCAAAATGAGAAAGATTCTTCAACTCCAACTGGAGATA	268		
DB	581	TGAATGATATAAATTTGGTAAAGCAGGTGAGACAACTTATTCACCTGGGCTTACATTTG	640		
QY	269	AAAAAGTTTATTCAGAAACAACTTTAACTACAGCTAATGGAATGAAGATTATGTAAGA	328		
DB	641	AAAGTGGTCTACAGAAAAAATTTGTTAAATTAGGGACAGTGATATTAAGATATAACTA	700		
QY	329	CAACTTTAAA---AATTTAGATGACAGGAAATATGCTATTATAGATTTAACTTATAATA	385		
DB	701	AAGCTCTTAACTTACTGTGTTCCTGGAAAGTAAAGCAACTGTTAAGTGTTCGAAAAA	760		
QY	386	ATGCTAAAACTGTG-----AAATTAAAGTAGTAGCAGCTAGTGAAA	427		
DB	761	CACCAAGTCCAGTGTTCAACAGTAATAACAAGCTTAGAATAATAAATGCTAAAGAG	820		
QY	428	AAACAGTAGTGTATCTAGTAGTGCAGAAAAATAGTGCAGAAAGATATAGCTGAAAAATG	487		
DB	821	AAACAAATAGATATTCAGCTAGTCTTAGTAAAAACAGCAGCAAGATTTAGCTAAAAAATG	880		
QY	488	TGTTTGAACACAAAGACTTAGAAAAATGCACATAAACTATAAATGCCTCAGATTTTCAGTA	547		
DB	881	TATTTAATAAACTGAATTTAAATACCTTTTATAAAGTATTAAATGGAGATGAAGCAGATA	940		
QY	548	AAACTGATAGTTACTATCAAGTAGTCTTTTATCCAAAAGGAAAGAGATTACAAAGTTTCT	607		
DB	941	CTAATGGAATTAATAGAAAGTTAGTGGAAATATCAAGTAGTCTTTTATCCAGAAAGNA	1000		
QY	608	CAACTTATAGAGCTACAAATTTAT---AATGAAGAACTGCATATGTTAATAACACAGTAA	664		
DB	1001	AAAGAGTTACAACTAAGAGTGTGCAGAGGCTTCAATTCCTGATGAAATTCACCAAGTTA	1060		
QY	665	TATTAACCTTAAATCTACTAGTAAAGTAATTTAAAGACTGCAGTAGAAGAGTTACAAA	724		
DB	1061	AAATTAACCTTAAGTCAGATAAAGAAAGACCTTAAAGATTATGTTGGATGATTTAAGAA	1120		
QY	725	AATTTGAATCTAGTTATTTCTAATACTACAACTTTTAGCTGGTGTGATGACAGAAATCAAAACAG	784		
DB	1121	CATATAAATGGAATTTCAATGCTATAGAGTAGCAGGAGAGATAGATAGAACTG	1180		
QY	785	CTATAGAGATAAGTAAAGAAATTTACAATAATGATGGCGAGAAATCAGATCAATTCAGCTG	844		
DB	1181	CAATAGCATTAAGTCAAAAATATTAATCACTCTGATGATGAAAA-----TGCTA	1228		
QY	845	ATGTTAAACAGAAATGTTAAAATGTTTATTAGTAGTGGTCAAAATGCACCTAGTAGATGGAT	904		
DB	1229	TATTTAGAGATTCAGTTGATAATGATGATATGGTTGGAGAAATGCAATGTTGATGGAC	1288		
QY	905	TAGTTGGCGCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACCTTCAAAAGATA	964		
DB	1289	TTGTAGCTTCTCTTAGCTCTGAAAGAAAGCTCCITTTATTAATTAATTCAAAAGATA	1348		
QY	965	AATTAGATTCGTCAGTAAATCTGAAATAAAGAGAGTTTGTAGCTTAAACCTCAACAG	1024		
DB	1349	AATTAGATTCAGCGGTAAAGCTGAAATAAAGAGAGTTATGAATATAAAGAGTACAAACAG	1408		
QY	1025	AAAGTA---ACAGAAAAACAGTTTATATAGTCTGGTGGAGTTAATAGTGTATCTAAAGAG	1081		

RESULT 4

US-10-068-870-6

; Sequence 6, Application US/10068870

; Publication No. US20030054009A1

; GENERAL INFORMATION:

; APPLICANT: The Provost, Fellows & Scholars of the College of the Holy and Undivided

; APPLICANT: Trinity of Queen Elizabeth, near Dublin

; TITLE OF INVENTION: C.difficile vaccine

; FILE REFERENCE: TRI002/C/WO

; CURRENT APPLICATION NUMBER: US/10/068,870

; CURRENT FILING DATE: 2002-02-11

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

DB	1409	GTATAAATACTTCAAGAAAGTTTATTTAGCTGGTGGAGTTAATCTTATCTATAATAAGAG	1468
QY	1082	TTGTAAACAGAAATTAGAATCAATGGGATTAAAGATTGAAAGATTCTCAGGTGATGATAGAT	1141
DB	1469	TAGAAAATGAATTAAGAATATGGGACTTAAAGTTACAAGATTAGCAGGAGATGATAGAT	1528
QY	1142	ATGAAACTTCTTTTAAAAATAGCAGGTGAAATAGGCTTATAGATTAATGATTAAGGCTTATG	1201
DB	1529	ATGAAACTTCTTAAAAATAGCTGATGAAGTAGGTCTTGAATGATAAAGCATTTGTTAG	1588
QY	1202	TTGTTGGGAACAGGATTAGCAGATGCCATGAGTATAGCTTCAAGTCTTCTACTTAATTAG	1261
DB	1589	TTGGAGGAACAGGATTAGCAGATGCCATGAGTATAGCTCCAGTTGCACTTCAATTAAGAA	1648
QY	1262	ATGTTAAATGGTGTCTAGATAGAACAAATGGACATGCTACTCCAATAGTTGTTGTAGATG	1321
DB	1649	ATGCTAAATGGTAAATGGATTAGCTGATGGTGTGCTACACCAATAGTTGTTGTAGATG	1708
QY	1322	GAAAAGCTGATAAATAATCTGATGATCTAGATAGTTTCTTAGGAAAGCGCTGATGATAGTA	1381
DB	1709	GAAAAGCTTAAACTATAAATGATGATGATAAAGATTCTTCTAGATGATTCACAAGTTGATA	1768
QY	1382	TAATAGGTGGATTTTCGAAGTGTATCTGAAAGATGGAAGAGCTATATCAGATGCTACTG	1441
DB	1769	TAATAGGTGGAGAAAACAGTGTATCTAAGATGTTGAAATGCAATAGATGCTGTACAG	1828
QY	1442	GTAAGAGGCGTTTACAAGAGTTAAAGCGCAGATAGACAAAGACACTAACTCTGAAAGTTATA	1501
DB	1829	GTAATCTCCAGATAGATATAGTGAGATGATAGACAGCAACTAATGCAAAAGTTATAA	1888
QY	1502	AAACATATATGCTAATGATGATGAAATAGCTTAAAGCTGCAGTTTTAGATTAAGATTGAT	1561
DB	1889	AGAACTCTTCTTATTAATCAAGATACTTAAATAATGATAAAGAGTAGTTAATTTCTTTG	1948
QY	1562	GTGCTTCAAGTAGTGCAGAGTATTTAATTTCTATGTAAGTAAAGATGATGATCTACAA	1621
DB	1949	TAGCTAAAGATGGTTCTTACTAAGAAAGATCAATAGTTGATGCTTTAGCAGCAGCTCCAG	2008
QY	1622	A-----AGAAAGTCAATTAGTTG	1639
DB	2009	TTGCAGCAAACTTTTGGTGTAACTCTTAATCTGATGGTAAAGCAGTAGATAAAGATGTA	2068
QY	1640	ATGCATTAAGCAGTAGGAGCTGTTGCTGGATATAAA-----CTTGCTCAGTTGTTAT	1690
DB	2069	AAATTAATTAAGTGTCTGATAATGATAAATAAATAAATAGTATCTCCAGCAGCTATAGTAT	2128
QY	1691	TAGCTACTGATTTCTTTATCTCTGATCAATCGTTGCTTATAGCAAAAGTTGTAGGAGAA	1750
DB	2129	TAGCTACTGATTTCTTTATCTTCAAGATCAAGTGTATCTATAGTAAAGTTCTTTGATAAG	2188
QY	1751	AATATTTTAAAGATTAAACACAAGTTGCTCAAGGAATAGCTAATTCAGTTATAAACAATA	1810
DB	2189	ATAATGGAGAAAACCTTAGTTCAAGTTGGTAAAGGTATAGCTACTTCACTTATAAACAAT	2248
QY	1811	TGAAAGATTATAGATATG	1830
DB	2249	TAAAGATTATTAAGTATG	2268

```

; SEQ ID NO 6
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Cl
US-10-068-870-6

```

Query Match	27.5%	Score 503.4	DB 14	Length 2217
Best Local Similarity	59.6%	Pred. No. 1.8e-69		
Matches 950	Conservative 0	Mismatches 616	Indels 27	Gaps 5
Qy	250	TCAACTCCAACTGCAGATAAAAAAGTTTATTTCAGAACAAACTTTTAACTACAGCTAATGGA	309	
Db	640	TCCTGTTCTACTACAGGCTTAACTTAACTGCTGATCACTGCAACACAGATGTAAT	699	
Qy	310	AATGAAGATTATGTAAGAACAATTTTAAAAAATTTAGATCGAGAGAAATATGCTATTATTA	369	
Db	700	ATTCTCATGTTATGAGTGCATTTAAATTTAAATTTGTTACTGATACGATTTAGTGGATTTCCCA	759	
Qy	370	GAATTAACATTATTAATATGCTTAAACCTGTTGAATTTAAAGTTAAAGCTAGTAGCAGCTAGTGA	429	
Db	760	GCTGGTTTCATCAGCTTCTACTCTTTAGAGCAAGTATTAAGATTAATGAATGCAAGAAGAA	819	
Qy	430	ACAGTAGTTGTATCTAGTGAATGCCAAAAATAGTGCAAAAATAGCTGAAAAATATGTC	489	
Db	820	TCATAGATTTGATTCAGGTTCCATAGAACACCTGAGATTTAGCTGAAAAATATGTA	879	
Qy	490	TTTGAAGACAAAGACTTAGAAAAATGCACTAAAAAATTAATGCTCAGATTTTC-----	543	
Db	880	TTTAAACCCAGAGATGTGAATAAACCTTATGAGGCACCTGATGATTTATATAAAGAAGGT	939	
Qy	544	AGTAAACTGATAGTACTTCACTCAAGTAGTCTTTTATCCAAAGGAAAGAGATTAACAGGT	603	
Db	940	ATAACAAGTAACTTATCACTCAAGATGGTGGAAAAATATCAAGTTGTTTATTTGCTCAA	999	
Qy	604	TTCTCAACTATAGAGCTACAAATTTAATGAAGAACTGTCATATGTTAATACACCACTA	663	
Db	1000	GGAAAGAGATTAACTACTTAAAGGAGCACTGGAACCTTTAGCAGATGAAATTTCTCTCTT	1059	
Qy	664	ATATTAACTCTAAAAATCTATAGTAAGATTAATTTAAAGACTGCAGTAGAGAGGTTACAA	723	
Db	1060	AAAGTAACAAATAAAGCAGATAAAGTAAAGACTTAAAGAAATATGTTGAAGATTTTAA	1119	
Qy	724	AAATTTGAATGCTAGTATTCTTAACTACTACAACTTTAGCTGGTGTGATGACAGAAATCA	783	
Db	1120	AATGCTAACAATGGATATTCAAATTCCTGTTGTTGTCAGAGTGAAGATAGATAGAA	1179	
Qy	784	GCTATPAGAGATAGTAAAGAAATATTAACAATTAATGATGCGAGAAATCAGATCATCT	843	
Db	1180	GCAATAGAGTTTAAAGTAGCAAAATCTATACTCTGATGATGACAA-----TGCA	1227	
Qy	844	GATGTTAAAGAGATGTTTAAAAATGTTGTTATTTAGTAGTGCGCAATGCACTAGTAGAG	903	
Db	1228	ATAACTAAAGATCCAGTTTAAACAATGTTGTTTGTGTCAGAGTGAAGATAGATAGAA	1287	
Qy	904	TTAGTTGGGCTCCTTTTAGCAGCAGAAAAAGATGCTCCACTATTATTATTAACCTTTCAA	963	
Db	1288	CTTTAGCTTCACCTTTAGCATCTGAAAAAAGAGCTCTTTTACTATTATTAACCTTCAG	1347	
Qy	964	AAATTTAGATTTCGTTCAGTAAAAATCTGAAATAAAGAGAGTTTGTAGACTTTAAAAA	1023	
Db	1348	AAATTTAGATTCAAAGTTTAAAGCTGAGTTTGAAGAGAGTAAATGAGTATCTTAAAGAA	1407	
Qy	1024	GAACTA---ACAGGAAAAACAGTTTATATAGCTGGTGAGTTTAAATAGTATCTTAAAGAA	1080	
Db	1408	GGTGTAAATATCTTCTTAAAAAAGTTTACTTAGCTGGTGAGTAACTCTTATATCTTAAAGT	1467	
Qy	1081	GTTTGTAAACAGAAATTAAGATCAATGGGATTTAAAGTTTGAAGATTTCTCAGGTCATAGA	1140	
Db	1468	GTAGAAATGAATTAAGATATGGAATTAAGATTAAGATTAATCAGGAGATGATAGA	1527	
Qy	1141	TATGAACTCTTTTAAAAATAGCAGGTGAATATAGCTTTAGATATAATATGATAGCTTATGTA	1200	
Db	1528	TATGAACTCTTTTATAGCTATAGCTAGTGAATAGCTCTTGTATTAATGAATAAGCTTTTGT	1587	

RESULT 5
US-10-239-610-4
; Sequence 4, Application US/10239610
; Publication No. US20040039165A1
; GENERAL INFORMATION:
; APPLICANT: Imperial College of Science Technology and Medicine
; TITLE OF INVENTION: CLOSTRIDIUM DIFFICILE POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: YOUZ 2 00089
; CURRENT APPLICATION NUMBER: US/10/239,610
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01305
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: GB 0007263.7
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-239-610-4

Query Match 26.4%; Score 483.2; DB 18; Length 2145;
Best Local Similarity 62.7%; Pred. No. 2.6e-66;
Matches 911; Conservative 0; Mismatches 433; Indels 108; Gaps 6;
QY 394 ACTGTTGAATTAAGTAGTAGCAGCTAGTGTCAAAAAACAGTAGTTGTATCTAGTGTATCCG 453

QY 751 ACACTTTAGCTGGTGAATGACAGACATACAAACAGCTATAGAGATAAGTAAAGAAATATTAC 810
DB |||||
QY 1144 GTAAACAGTACGAGAGAAGATAGAAATGAAGAACTGCTATAGAAATTAAGTAAATATTAT 1203
DB |||||
QY 811 AATAAGTATGCGCAGAGAAATCAGATCAATTCAGCTGATGTTAAAGAGAAATGTTAAAAATGTT 870
DB |||||
QY 1204 AATCTGATGATAAATAATCAATACTGATGATGACAGTTAAT-----AATATA 1251
DB |||||
QY 871 GTATTAGTAGGTCAGAAATGCACTAGTAGATGGAATTTAGTTCGGCTCTTTAGCAGCAGAA 930
DB |||||
QY 1252 GTATTAGTAGGATCTACATCTATAGTTGATGGTCTTGTTCATCACCATTAGCTTCAGAA 1311
DB |||||
QY 931 AAGATGCTCCATATTATTAATCTCAAAAGATAAATAATAGATTCGTCAGTAAATCTGAA 990
DB |||||
QY 1312 AAAACAGCTCCATTTATTATTAATCTCAAAAGATAAATAATAGATTCATCAGTAAATCTGAG 1371
DB |||||
QY 991 AATAAGAGAGTTTTAGACTTTAAAACTTCAACAGAGTA---ACAGGAAAAACAGTTTAT 1047
DB |||||
QY 1372 AATAAGAGAGTATGAATTTAAGAGTGATGATGATTAATAATCTCTAAAAAGTTTAT 1431
DB |||||
QY 1048 ATAGCTGGTAGGATTAATAGTGATCTAAAGAGTTGTTAAGAGATTTAGAAATCAATGGGA 1107
DB |||||
QY 1432 TTAGCTGGTAGGATTAATCTTATATCTAAAGATGTAGAGATGAATTTGAAAAATATGGGC 1491
DB |||||
QY 1108 TTAAGAGTTGAAAGATTTCTCAGGTGATGATAGATATGAAATCTTTTAAAAATAGCAGGT 1167
DB |||||
QY 1492 CTTAAAGTTTACTAGATTTATCAGGAGAAAGACAGATACGAACTCTTTAGCAATAGCTGAT 1551
DB |||||
QY 1168 GAAATAGGCTTAGATAATGATAGGCTTATGTTAGTTCGGTGAACAGGATTTAGCAGATGCC 1227
DB |||||
QY 1552 GAAATAGGCTTGTATATGATTAAGCAATTTGTTAGTTCGGTGTCTGGATTCGCAGATGCT 1611
DB |||||
QY 1228 ATGAGTATAGCTTCAGTTGCTTCTAATAATTTAGATGGTAATGTTGTAGATAGAAACA 1287
DB |||||
QY 1612 ATGAGTATAGCTTCAGTTGCTTCT-----CAACTTAAA 1644
DB |||||
QY 1288 AATGACATGCTACTCCAAATAGTTGTTAGATGGAAGCTGATAAATATCTGATGAC 1347
DB |||||
QY 1645 GATGAGATGCTACTCCAAATAGTTGTTAGATGGAAGCAAAAGAAATAGTGTATGAT 1704
DB |||||
QY 1348 TTAGATAGTTCTTTAGGAAGCGCTGATGATAGATATAATAGTGGATTTGCAAGTGTATCT 1407
DB |||||
QY 1705 GCTAAGAGTTCTTTAGGAACTTCTGATGTTGATATAATAGTGGAAATAAGCTATCT 1764
DB |||||
QY 1408 GAAAGATGGAAGAGCTATATCAGATGCTACTCGTTAAAGCGTTTACAAGATTTAAAGGC 1467
DB |||||
QY 1765 AAAGAGATTGAAGAGTCAATAGATAGTGCACACTGGAAAACTCCAGATAGAATAAGTGA 1824
DB |||||
QY 1468 GACATAGACAAGACACTTAATCTGAAAGTTATATAAACAATATTATGCTAATGATCTGAA 1527
DB |||||
QY 1825 GATGACAGACAAGCAACTTAATGCTGAAAGTTTAAAA----- 1860
DB |||||
QY 1528 ATAGCTAAAGCTGCAGTTTTAGATAAAGATTCAGTGCTTCAAGTGTGATGATGAGGAGTA 1587
DB |||||
QY 1861 -----GAAATGATTAATTTCAAGATGGTGAAGTT 1890
DB |||||
QY 1588 TTTAATTTCTATGCTAGCTAAAGATGGATCTCAAAAGAGATCAATATTAGTTGATGCAATTA 1647
DB |||||
QY 1891 GTGAATTTACTTTGTTGCAAAAGATGGTTCTACTAAAGAGATCAATATTAGTAGATGCAATTA 1950
DB |||||
QY 1648 GCAGTAGAGCTGTTGCTGG-----ATATAAACTTCTCCAGTTGTTATTAGCTACT 1698
DB |||||
QY 1951 GCAGCAGCACCAGTGTAGATTTAAGGAGTCTCCAGCTTCAATCATACTAGCTACT 2010
DB |||||
QY 1699 GATCTTTTATCTTCTGATCAATCGGTTGCTTAAGCAAGTTGTAGGAGAAAAATATTCT 1758
DB |||||
QY 2011 GATATCTTTATCTTCTGACCAAAATGATGATGTTAAGTAAAGCAGTTCTTAAAGATGGTGA 2070
DB |||||
QY 1759 AAAGATTTAAACAAGATGGTCAAGGAAATAGCTAATTCAGTTTATAAACAATAAGAT 1818
DB |||||
QY 2071 ACTAATCTTAGTTCAAGTAGGTAAGGATAGCTTCTTCAGTTTATAAACAATAAGAT 2130
DB |||||

QY 1819 TTATTAGATATG 1830
DB |||||
QY 2131 TTATTAGATATG 2142
DB |||||

RESULT 7

US-10-068-870-1

; Sequence 1, Application US/10068870

; Publication No. US20030054009A1

; GENERAL INFORMATION:

; APPLICANT: The Provost, Fellows & Scholars of the College of the Holy and Undivided

; TITLE OF INVENTION: Trinity of Queen Elizabeth, near Dublin

; FILE REFERENCE: TRI002/C/WO

; CURRENT APPLICATION NUMBER: US/10/068,870

; CURRENT FILING DATE: 2002-02-11

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: Patent version 3.1

; SEQ ID NO 1

; LENGTH: 2157

; TYPE: DNA

; ORGANISM: Clostridium difficile

US-10-068-870-1

Query Match 25.8%; Score 471.8; DB 14; Length 2157;

Best Local Similarity 61.5%; Pred. No. 1.6e-64;

Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;

QY 296 CTACAGCTAATGGAATGAAGATTATGTAAGACAACTTTAAAAATTTAGATGAGGAG 355
DB |||||
QY 728 CAACACCTAGTGTGCTAGCTGTAAGTGGTTTGTAACTAAAGATGATCTGATTTAGCAA 787
DB |||||
QY 356 AATATCTCTATTATAGATTTAACTTATAATAATGCTAAAACTGTTGAAATTTAAAGTAGTAG 415
DB |||||
QY 788 AATCAGTACTATTAATGTAAGGTTTATAATGCAAAAGAGAAATCAATGATATAGATG 847
DB |||||
QY 416 CAGCTAGTGAATAAACAAGTAGTTGTATCTAGTATGCGAAAAATAGTGCAAAAAGATATAG 475
DB |||||
QY 848 CAGCTCATATACATCAGCTG-----AAAAATTAGCTAAAGATATGTTATTTGATCCAG 901
DB |||||
QY 476 CTGAAAAATATGTTGTTGAAGACAAAGACTTGAAGAAATGCACTTAAAACTTATAAATGCGCT 535
DB |||||
QY 902 ATGAAATTTCTGAAGCATATAGGCAATAGTATGATCAAAATGATGGTATAGATCTA 961
DB |||||
QY 536 CAGATTTCACTAAAACTGATAGTACTATCAAGTAGTTCTTTATCCAAAAAGGAAAGAGAT 595
DB |||||
QY 962 ATTTAGTTCAGTTAGTTAATGGAATATCAAGTATTTTATCCAGAGGTAAAGAT 1021
DB |||||
QY 596 TACAAGGTTTCTCAACTTATAGAGCTACAAATTTAATGAAGGAACCTGCATATGGTAATA 655
DB |||||
QY 1022 TAGAACTAAATCAGCA-----AATGATACAATAGCTAGTCAAGATA 1063
DB |||||
QY 656 CACAGTAAATTTAACTCTAAATCTACTAGTAAGAGTAATTTAAAGACTGCGAGTAGAG 715
DB |||||
QY 1064 CACCAGCTAAGTAGTTTATAAAGCTAATAAATTTAAAGATTTAAAAAGATTTATGTAGATG 1123
DB |||||
QY 716 AGTTACAAAAATTTGAATGCTAGTTATTCTAATACTACAATTTAGCTGGTGTGATGACAGAA 775
DB |||||
QY 1124 ATTTAAAAACATATAATATCTTTTCAATGTTGTAACAGTAGCAGGAGAGATAGAA 1183
DB |||||
QY 776 TACAAAACAGCTATAGAGATAAGTAAAGAAATTTACAATATATGATGGCGAGAAATTCAGATC 835
DB |||||
QY 1184 TAGAACTGCTATAGAAATTAAGTAGTAAATATTATATATCTCGAT-----GATA 1231
DB |||||
QY 836 ATTCAGCTGATGTTAAAGAGAAATTTAAAAATGTTGTTATTAGTAGGTGCAATGCACTAG 895
DB |||||
QY 1232 AAAATGCAATAAATCTGATATAAGCAGTTAATGATATAGTATTAGTTGGATCTACATCTATAG 1291
DB |||||
QY 896 TAGATGATTTAGTTGCGGCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTT 955
DB |||||
QY 1292 TTGATGCTGTTGTCATCACCATTAGCTTCCAGAAAAAACAGCTCCCATTTATTAACTT 1351
DB |||||
QY 956 CAAAAGATAAATTAGATTGCTGCTAGTAAAAATCTGAAATTAAGAGAGTTTTAGACTTAAAAA 1015
DB |||||

```

1352 DB CAAAAGTAAATAGATTCATCACTAAATCTGAAATAAAGAGATTGAACTTAAAGA 1411
1016 QY CTTCAACAGAGTA---ACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAAATAGTGAT 1072
1412 DB GTGACACTGGTATAAATACCTCTAAAAAAGTTTATTTAGCTGGTGGAGTTAAATCTATAT 1471
1073 QY CTAAGAAGTTGTAAACAGATTAGATCAATGGGATTAAGGTTAAAGATTCACAGGTG 1132
1472 DB CTAAGAAGTTGTAGAAAATGAATTGAAAACAATGGGCTTAAAGTTACTAGATTATCAGGAG 1531
1133 QY ATGATAGATATGAACTCTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGG 1192
1532 DB AAGACAGATACGAACTCTTTAGCAATAGCTGATGAAATAGTCTTGATATGATAAG 1591
1193 QY CTTATGATGTTGGTGAACAGGATTAGCAGATGCCATGATGATAGCTTCAGTTGCTTCTA 1252
1592 DB CATTTGATGTTGGTGGTACTGGATTAGCAGATGCTATGATATAGCTCCAGTTGCTTCT- 1650
1253 QY CTAATATAGTGTATGTTGTTAGTATAGAACAAATGGACATGCTACTCCATAGTTG 1312
1651 DB -----CAACTTAAAGATGGAGATGCTACTCCAAATAGTAG 1684
1313 QY TTGTAGATGAAAAGCTGATAAATATCTGATGACTTAGATAGTTTCTTAGGAAGCGCTG 1372
1685 DB TTGTAGATGGAAGAAGCAAGAATAAAGTATGATGCTCTAAGAGTTTCTTAGGAATCTCTG 1744
1373 QY ATGTAGATATAATAGGTGGATTTGCAAGTGTATCTGAAAAGATGGAAGAGCTATATCAG 1432
1745 DB ATGTTGATATAATAGGTGGAAAAAATAGCGTATCTAAAGAGATTGAAGAGTCAATAGATA 1804
1433 QY ATGCTACTGTTAAGCGCTTACAAGAGTTAAAGGCGACGATAGACAAGACACTAACTCTG 1492
1805 DB GTGCAACTGGAAAAAATCTCAGATAGAAATAAGTGGAGATGATAGACAAGCAACTAAATGCTG 1864
1493 QY AAGTTATAAAAAACATATTATGCTAAATGATACATACTGAAATAGCTAAAGCTGCAGTTTTAGATA 1552
1865 DB AAGTTTAAAGAAGATGATTATTC----- 1890
1553 QY AAGATTCAAGTGCTTCAAGTAGTGATGCGAGGAGTATTTAATTTCTATGCTAGCTAAAGATG 1612
1891 DB -----ACAGATGGTGAAGTTGTGAATTACTTTGTTGCAAAAAGATG 1930
1613 QY GATCTCAAAAAGAGATCAATTAGTTGATGATGATGATGATGATGATGATGATGATGATG 1667
1931 DB GTTCTACTAAAGAAGATCAATTAGTAGATGCTCTAGCAGCAGCAGCAGCAGCAGTAGAT 1990
1668 QY ----ATATAAACTTGCTCCAGTTGTATTAGCTACTGATTTCTTTATCTTCTGATCAATCGG 1723
1991 DB TTAAAGAGTCTCAGCTCCATCATCTAGCTACTGATTTATTTATCTTCTGACCAAAATG 2050
1724 QY TTGCTATAAGCAAAAGTTGTAGGAGAAAAATATTTCTAAAGATTTTAAACAAGTTGGTCAAG 1783
2051 DB TAGCTGTAAAGTAAAGCAGTTCTCTAAAGATGGTGGAACTAACTTAGTTCAAGTAGGTAAAG 2110
1784 QY GAATAGCTAATTCAGTTTATAACAAAATGAAAGATTTATAGTATG 1830
2111 DB GTATAGCTTCTTCAGTTTATAAACAANAATGAAGAATTTATAGATATG 2157

```

```

RESULT 8
US-10-068-870-3
; Sequence 3, Application US/10068870
; Publication No. US20030054009A1
; GENERAL INFORMATION:
; APPLICANT: The Provost, Fellows & Scholars of the College of the Holy and Undivided
; TITLE OF INVENTION: C.difficile vaccine
; FILE REFERENCE: TRI002/C/WO
; CURRENT APPLICATION NUMBER: US/10/068, 870
; CURRENT FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 3
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-068-870-3

Query Match      25.8%; Score 471.8; DB 14; Length 2158;
Best Local Similarity 61.5%; Pred. No. 1.6e-64;
Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;

QY 296 CTCAGCTAAATGGAATGAAGATTATGTAAAGACAACCTTTTAAAAATTTAGATGCGAGAG 355
DB 728 CAACACCTAGTCTGCTAGCTGTAAGTGGTTTTGTAACTAAAGATGATCTAGTTTAGCAA 787
QY 356 AATATGCTATTTATAGATTAACTTATATAATGCTTAAAACTGTTGAAATTTAAAGTAGTAG 415
DB 788 AATCAGGTACTATAAATGTAAAGATTATAAATGCAAAAAGAAATCAATTTGATATAGATG 847
QY 416 CAGCTAGTGAAAAAACAGTAGTTGTATCTAGTAGTGCGAAAAAATAGTGCAAAAGATATAG 475
DB 848 CAAGCTCATATACATCAGCTG-----AAAATTTAGCTTAAAGATATGATTTTGATCCAG 901
QY 476 CTGAAAAATATGTTGTTGAAGACAAGAGCTTAGAAAAATGCCTAAAAAACTATAAAATGCCCT 535
DB 902 ATGAAATTTCTGAAGCATATAGGCAATAGTAGCATTACAAAATGATGCTATAGAGTCTA 961
QY 536 CAGATTTCAAGTAAAACTGATAGTTACTATCAAGTAGTTCTTTTATCCAAAAGGAAAGAGAT 595
DB 962 ACTTAGTTTCAGTTAGTTAATGGAATAATATCAAGTGAATTTTTTATCCAGAAGGTAAAGAT 1021
QY 596 TACAAGGTTTCTCAACTTATAGAGCTACAAAATTTATAATGAAGAACTGCATATGGTAAATA 655
DB 1022 TAGAACTAAATCAGCA-----AATGATACAATAGCTAGTCAAGATA 1063
QY 656 CACCAGTAAATATAAATCTCTAAATCTACTAGTAGAAGAGTAAATTTAAAGACTGCGAGTAGAG 715
DB 1064 CACCAGCTAAAGTAGTTATAAAGCTAATAAATTTAAAGATTTTAAAGAGTTATGTAGATG 1123
QY 716 AGTTACAAAATTTGAATGCTAGTTATTTCTAATACTACAACTTTAGCTGGTGTATGACAGAA 775
DB 1124 ATTTAAAAACATATAATAATACTTATTTCAAAATGTTTGAACAGTAGCAGGAGAAAGATAGAA 1183
QY 776 TACAACAGCTATAGAGATAAGTAAAGAAATATTACAATAATGATGCGAGAAATCAGATC 835
DB 1184 TAGAACTGCTATAGAAATTAAGTAGTAATAATTTAATTTCTGAT-----GATA 1231
QY 836 ATTCAGCTGATGTTAAAGAGAAATGTTTAAAAATGTTGTATTTAGTAGTGCAAAATGCACTAG 895
DB 1232 AAAATGCAATRACTGATAAGCAGTTAATGATATAGTATTAGTTGGATCTACATCTATAG 1291
QY 896 TAGATGGATTAGTTGGGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTT 955
DB 1292 TTGATGGTCTTGTGTCATCACCATTAGCTTCCAGAAAAAACAAGCTCCCAATTTATTAACTTT 1351
QY 956 CAAAAGATAAAATTAGATTGCTCAGTAAATCTGAAATAAAGAGAGTTTTTATAGACTTTAAAAA 1015
DB 1352 CAAAAGATAAAATTAGATTGCTCAGTAAATCTGAAATAAAGAGAGTTTATGAATTTAAAGA 1411
QY 1016 CTTCAACAGAGATA---ACAGGAAAAACAAGTTTTATATAGCTGGTGGAGTTAATAGTGAT 1072
DB 1412 GTGACACTGGTATAAATCTCTTAAAAAAGTTTTTATTTAGCTGGTGGAGTTAATTTCTATAT 1471
QY 1073 CTAAGAAGTTGTATAACAGAAATTTAGAAATCAATGGGATTTAAAGTTGAAAGATTTCTCAGGTG 1132
DB 1472 CTAAGAAGTTGTAGAAAATGAATTTGAAAAACAATGGGCTTTAAAGTTACTAGATTATCAGGAG 1531
QY 1133 ATGATAGATATCAAACTCTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGG 1192
DB 1532 AAGACAGATACGAACTCTTTAGCAATAGCTGATGAAATAGTCTTGATATGATAAG 1591
QY 1193 CTTATGATGTTGGTGAACAGGATTAGCAGATGCCATGATGATAGCTTCAGTTGCTTCTA 1252
DB 1592 CATTTGATGTTGGTGGTACTGGATTAGCAGATGCTATGATATAGCTCCAGTTGCTTCT- 1650

```

QY 1253 CTAAATTAGATGGTAATGGTGTCTAGATAGAACAAATGACATGCTACTCCCAATAGTTG 1312
 Db 1651 -----CAACTTAAAGATGGAGATGCTACTCCCAATAGTAG 1684
 QY 1313 TTGTAGATGGAAGCTGATAAAATATCTGATGACTTTAGATAGTTCTTAGGAAGCGCTG 1372
 Db 1685 TTGTAGATGGAAGCAAAAGAAATAGTGTATGATGCTAAGAGTCTTCTTAGGAACCTCTG 1744
 QY 1373 ATGTAGATATAATAGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATCAG 1432
 Db 1745 ATGTTGATATAATAGTGGGAAAAATAGCGTATCTAAAGAGATTTGAAGAGCTCAATAGATA 1804
 QY 1433 ATGCTACTGTTAAAGCGTTTACAGAGTTAAAGCGAGCATAGACAGACACTAACTCTG 1492
 Db 1805 GTGCAACTGGAAAAAACTCCAGATAGAAATAGTGGAGATGATAGCAAGCAACTANTGCTG 1864
 QY 1493 AAGTTTAAAAACAATATATGCTAATGATATCTGAAATAGCTAAAGCTGCAGTTTTAGATA 1552
 Db 1865 AAGTTTAAAAAGATGATTAATTC----- 1890
 QY 1553 AAGATTCAGGTGCTCAAGTAGTATGATGATGATGATGATGATGATGATGATGATGATG 1612
 Db 1891 -----ACAGATGCTGAAGTTGTGAATTAATCTTTGTTGCAAAAGATG 1930
 QY 1613 GATCTACAAAGAGATCAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1667
 Db 1931 GTTCTACTAAGAAGATCAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1990
 QY 1668 ----ATATAAACTTGCTCCAGTGTCTATTAAGTGTGATGATGATGATGATGATGATGATG 1723
 Db 1991 TTAAGGAGTCTCCAGTCTCAATCATAGTACTGATGATGATGATGATGATGATGATGATG 2050
 QY 1724 TTGCTATAAGCAAAAGTTGTAGGAGAAATATCTTAAAGATTTAAACACAGTTGGTCAAG 1783
 Db 2051 TAGCTGAAGTAAAGCAGTCTCTAAAGATGTTGGAAGTAACTTAAGTTCAAGTAGGTAAAG 2110
 QY 1784 GAATAGCTAATTCAGTTATTAACAAATGAAGATTTTATTAGATATG 1830
 Db 2111 GTATAGCTTCTCAGTTATTAACAAATGAAGATTTTATTAGATATG 2157

RESULT 9

US-10-068-870-8
 ; Sequence 8, Application US/10068870
 ; Publication No. US20030054009A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Provost, Fellows & Scholars of the College of the Holy and Undivided
 ; APPLICANT: Trinity of Queen Elizabeth, near Dublin
 ; TITLE OF INVENTION: C.difficile vaccine
 ; FILE REFERENCE: TRI002/C/WO
 ; CURRENT APPLICATION NUMBER: US/10/068,870
 ; CURRENT FILING DATE: 2002-02-11
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 2158
 ; TYPE: DNA
 ; ORGANISM: Clostridium difficile
 US-10-068-870-8

Query Match 25.8%; Score 471.8; DB 14; Length 2158;
 Best Local Similarity 61.5%; Pred. No. 1.6e-64;
 Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;
 QY 296 CTACAGCTAATGGAATGAAGATTATGTAAGACAACTTTTAAAAATTTAGATGAGCAG 355
 Db 728 CAACACCTAGTCTGTAGCTGTAAGTGGTTTTGTTAACTTAAGATGATGATGATGATGATG 787
 QY 356 AATATGCTATTATAGATTTAACTTATAATTAATGCTTAAATCTGTTGGAATTTAAAGTAG 415
 Db 788 AATCAGGTACTATAATGTAAGAGTTATAAATGCCAAAGAAAGAAATCAATTTGATATGATG 847

QY 416 CAGCTAGTGAAGAAAAACAGTAGTTGTATCTAGTGTAGCGAAAAAATAGTGCACAAAGATATAG 475
 Db 848 CAAGCTCATATACATCAGCTG-----AAAAATTAGCTAAAGACATGTATTTTGTATCCAG 901
 QY 476 CTGAAAAATATGTTGTTGAAGACAAAGACTTAAAGAAATGCATTAATAAATGCATATAAATGCCT 535
 Db 902 ATGAAATTTCTGAAGCATATATAGGCAATAGTATGATTTACAAATATGATGGTATAGAGTCTA 961
 QY 536 CAGATTTCAAGTAAAACTGATAGTTACTATCAAGTAGTCTTCTTTATCCAAAGGAAAGAGAT 595
 Db 962 ATTTAGTTCAGTTAGTTAATGGAATAATCAAGTAGTATTTTATCCAGAGGTTAAAGAT 1021
 QY 596 TACAAGGTTTCTCAACTTATAGAGCTCAAAATATTAATGAAGAACTGCATATGGTAATA 655
 Db 1022 TAGAAAACTAAATCAGCA-----AATGATACAAATAGCTAGTCAAGATA 1063
 QY 656 CACCACTAATATTAACCTCTAAATCTACTAGTAAGAGATTAATTTAAAGAGCTGCAGTAGTAAG 715
 Db 1064 CACCACTAAGTAGTTATAAAAGCTAATAAATTAAGATTTTAAAGATTTATGATAGTG 1123
 QY 716 AGTTACAAAAATTTGAATGCTAGTTATTTCTAATACTACAATTTAGCTGGTGTGATGACAGAA 775
 Db 1124 ATTTAAACATATATATATATCTTATTTCAATCTGTTAAACAGTAGCAGGAGAAAGATAGAA 1183
 QY 776 TACAAACAGCTATAGAGATAAGTAAAGATATTTACAAATATGATGGCGAGAAATCAGATC 835
 Db 1184 TAGAACTGCTATAGAAATTAAGTAGTAAATATTTATTAATCTGAT-----GATA 1231
 QY 836 ATTCAGCTGATGTTAAAGAGAAATGTTAAAAATCTGTTATTTAGTGGTGCAGATGCACATG 895
 Db 1232 AAAATGCAATPAACCTGATAAAGCAGTTAATGATATAGTATTTAGTTGGATCTACATCTATAG 1291
 QY 896 TAGATGATAGTTGCGGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTTATTAATCTT 955
 Db 1292 TTGATGCTCTGTTGCTACCATAGCTTTCAGAAAAACAGCTCCATTTATTTAACTT 1351
 QY 956 CAAAGAGATAAATAGATTCCTGATGATAAATCTGAAATTAAGAGAGTTTATAGACTTAAAA 1015
 Db 1352 CAAAGAGATAAATAGATTCATCACTAGTAAATCTGAAATTAAGAGAGTTTATGAATTAAGA 1411
 QY 1016 CTTCAACAGAGTA---ACAGGAAAAACAGTTTATATAGCTGTTGGAGTTTAAATAGTGTAT 1072
 Db 1412 GTGACACTGGTATAAATCTTCTAAAGAAAGTTTATTTAGCTGGTGGAGTTTAAATCTATAT 1471
 QY 1073 CTAAGAAAGTTGTAACAGAAATTAAGATCAATGGAATTTAAAGTTTGAAGATTTCTCAGGTG 1132
 Db 1472 CTAAGATGTAGAAAAATGAATTTGAAAAACATGGCTTTAAAGTTTACTAGATTATCAGAG 1531
 QY 1133 ATGATAGATATGAACCTTCTTTAAAAATAGCAGGTGAATAGGCTTATAGATTAATGATAAG 1192
 Db 1532 AAGACAGATACGAAACTTCTTTAGCAATAGCTGATGAATAGGCTCTTTGATTAATGATAAG 1591
 QY 1193 CTTATGATGTTGGTGGAAACAGGATTAGCAGATGCTAGTATAGCTTCAAGTCTCTTCTTA 1252
 Db 1592 CATTTGATGTTGGTGGTACTGGATTAGCAGATGCTATGAGTATAGCTCCAGTTGCTTCT- 1650
 QY 1253 CTAATATAGTGGTAAATGGTGTCTAGATAGAAACAAATGGAACATGCTACTCTCAATAGTTG 1312
 Db 1651 -----CAACTTAAAGATGGAGATGCTACTCTCAATAGTAG 1684
 QY 1313 TTGTAGATGGAAGAGCTGATAAAATATCTGATGACTTAGATAGTTTCTTAGGAAGCGGTG 1372
 Db 1685 TTGTAGATGGAAGAGCAAAAGAAATAGTGTATGATGCTTAAGAGTTTCTTAGGAACCTTCTG 1744
 QY 1373 ATGTAGATATATAGTGGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATCAG 1432
 Db 1745 ATGTTGATATAATAGTGGAAAAAATAGCGTATCTTAAAGAGATTTGAAGAGCTCAATAGATA 1804
 QY 1433 ATGCTACTGTTAAAGCGGTTTACAAGAGTTTAAAGCGAGCATAGACAGACACTAACTCTG 1492
 Db 1805 GTGCAACTGGAAAAAACTCCAGATAGAAATAGTGGAGATGATAGACACAGCAACTAATGCTG 1864
 QY 1493 AAGTTATAAAAAACATATTTATGCTAATGATGATACTGAAATAGCTTAAAGCTGCAAGTTTTAGATA 1552

```

1865 AGTTTTTAAAGAGATGATTTC----- 1890
1553 AAGATTGAGTCTTCAAGTAGTGATGCGAGGAGTATTTAAATTTCTATGTAGCTTAAAGATG 1612
1891 -----ACAGATGGTGAAGTGTGAATTACTTTGTTGCAAAAGATG 1930
1613 GATCTACAAAGAGATCAATTAGTTGATGCGATTTAGCAGTAGGAGCTGTTGCTGG----- 1667
1931 GTTCTACTAAGAGATCAATTAGTAGATGCCCTTAGCAGCAGCAACCAATAGCAGGTAGAT 1990
1668 ----ATATAAACTTGGCTCCAGTTGTTATAGTACTGATCTTCTTCTGATCAATCGS 1723
1991 TTAAGAGCTCCAGCTCCCAATCATACTAGTACTGATCTTTATCTTCTGACCAAAATG 2050
1724 TTGCTATAAGCAAGTTGTAGGAGAAAAATATTCTAAAGATTTTAAACAAGTTGGTCAAG 1783
2051 TAGCTGTAAGTAAAGCAGTTCTTAAAGATGGTGGAACTAACTTTAGTTCAAGTAGGTAAG 2110
1784 GAATAGCTAATTGAGTTATTAACAAAATGAAGATTTTATAGATATG 1830
2111 GTATAGCTTCTTCAGTTATATAACAAAATGAAGATTTTATAGATATG 2157

RESULT 10
US-10-239-610-5
; Sequence 5, Application US/10239610
; Publication No. US20040039165A1
; GENERAL INFORMATION:
; APPLICANT: Imperial College of Science Technology and Medicine
; TITLE OF INVENTION: CLOSTRIDIUM DIFFICILE POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: YOUZ 2 00089
; CURRENT APPLICATION NUMBER: US/10/239, 610
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01305
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: GB 0007263.7
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-239-610-5

Query Match 25.8%; Score 471.8; DB 18; Length 2160;
Best Local Similarity 61.5%; Pred. No. 1.6e-64;
Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;

QY 296 CTACAGCTAATGGAATGAAGATTATGTAAAGACAACTTTTAAAGAAATTTAGATGCGAGGAG 355
DB 728 CAACACCTAGTCTGTAGCTGTGAAGTGGTTTGTAACTAAAGATGATGATCTGATTTAGCAA 787
QY 356 AATATGCTATTATAGATTTTAACTTATATAAATCTTAAACTGTTTGAATTTAAAGTAGTAG 415
DB 788 AATCAGGCTACTATAATGTGAAGAGTTTATAAATGCAAAAGAAAGAAATCAATTTGATATAGATG 847
QY 416 CAGCTAGTGAAGAAACAGATGTTGTATCTAGTATGCGAAATAATAGTGCAGAAAGATATAG 475
DB 848 CAAGCTCATATACATCAGCTG-----AAAATTTAGCTAAAGATATGATTTTGTATCCAG 901
QY 476 CTGAAAAATATGTTGTTGAAGACAAAGACTTAGAAAAATGCACTTAAATAATATATAATGCTT 535
DB 902 ATGAAATTTCTGAAGCATATATAGGCAATAGTACATTAACAAATGATGGTATAGAGTCTA 961
QY 536 CAGATTTTCAGTAAAACTGATAGTTACTATCAAGTAGTCTTTTATCCAAAAGGAAGAGAT 595
DB 962 ATTTAGTTCAGTTAGTTTAAATGGAATAATATCAAGTGAATTTTTTATCCAGAAAGGTAAAGAT 1021
QY 596 TACAAGGTTTCTCAACTTATAGAGCTACAAATTTATTAATGAAGGAACTGCTATATGGTAATA 655
DB 1022 TAGAACTAAATCAGCA-----AATGATACAATAGCTAGTCAAGATA 1063

```

```

QY 656 CACCAGTAATATTAACCTCTAAAAATCTACTAGTAAGAGTAATTTTAAAGACTGCGAGTAGAAG 715
DB 1064 CACCAGCTAAAGTAGTTTAAAGAGCTAAATAAAATTAAGAGATTTTAAAGAGATTTAGTAGATG 1123
QY 716 AGTTACAAAATAATGGAATGCTAGTATTCTTAATCTACCAACTTTAGCTGGTGTATGACAGAA 775
DB 1124 ATTTAAACAATATTAATTAATCTTATTCAATGTTGTTAAACAGTAGCAGGAGAAAGATAGAA 1183
QY 776 TACAACAGCTATAGAGATAAGTAAAGAAATATTACAATAATGATGCGCGAGAAATCAGATC 835
DB 1184 TAGAACTGCTATAGAAATTAAGTAGTAAATATTATAATTTCTGAT-----GATA 1231
QY 836 ATTACGCTGATGTTTAAAGAGAAATGTTAAAAATGTTGTTATTAGTAGTGCAAAATGCACTAG 895
DB 1232 AAAATGCAATAACTGATAAAGCAGTTAAATGATATAGTATTAGTTGGATCTACATCTATAG 1291
QY 896 TAGATGGATTTAGTTGCGGCTCCTTTAGCAGCAGAGAAAAAGATGCTCCACTATTATTAACTT 955
DB 1292 TTGATGGTCTTGTGTCATCACCATTAGCTTTCAGAAAAAACAGCTCCATTATTATTAACTT 1351
QY 956 CAAAAGATAAAATTTAGATTTCGTCAGTAAATCTGAAATTAAGAGAGAGTTTATGAGCTTTAAAA 1015
DB 1352 CAAAAGATAAAATTTAGATTTCATCAGTAAATCTGAAATTAAGAGAGAGTTTATGAACTTAAAGA 1411
QY 1016 CTTCAACAGAGTA---ACAGGAAAAACAGTTTATATAGCTGCTGGTAGGTTAATAGTAT 1072
DB 1412 GTGACACTGGTATAAATACTTCTTAAAAAGATTTTATTAGCTGGTGGAGTTAATTTCTATAT 1471
QY 1073 CTAAGAAGAGTTCTTAACAGAAATTTAGAAATCAATGGGATTTAAAGCTTGAAGATTTCTCAGG 1132
DB 1472 CTAAGAAGAGTTAGAAAATGAAATGAAACATGGGCTTTAAAGTTACTAGATTATCAGGAG 1531
QY 1133 ATGATAGATATGAAACTTCTTTTAAAAATAGCAGAGTGAAATAGGCTTTAGATAATATGATAAG 1192
DB 1532 AAGACAGATACGAAACTTCTTTAGCAATAGCTGTAGTAAATAGCTCTTGATATATGATAAG 1591
QY 1193 CTTATGTTAGTTGGTGAACAGAGATTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTA 1252
DB 1592 CATTTGTTAGTTGGTGTACTGGAATAGCAGATGCTATGAGTATAGCTCCAGTTGCTTCT- 1650
QY 1253 CTAATTTAGATGTTGTTGTAGATAGAAACAAATGGACATGCTACTCCAAATAGTTG 1312
DB 1651 -----CAACTTAAAGATGGAGATGCTACTCCAAATAGTAG 1684
QY 1313 TTGTAGATGAAAAAGCTGATAAAAAATCTGATGACTTATAGATAGTTTCTTTAGAAAGCGCTG 1372
DB 1685 TTGTAGATGAAAAAGCAAAAGAAATAAGTGATGATGCTTAAAGAGTTTCTTAGGAACTTCTG 1744
QY 1373 ATGTAGATATATAGTGGATTGTCAGATGTTATCTGAAAAGATGGAAGAGCTATATACAG 1432
DB 1745 ATGTTGATATATAAGTGGAAAAAATAGCGTATCTTAAAGAGATTGAAGAGTCAATAGATA 1804
QY 1433 ATGCTACTGGTAAAGCGTTTACAAGAGTTTAAAGCGCAGATAGACAGACACTAACTCTG 1492
DB 1805 GTGCAACTGGAAGAAACTCCAGATAGATAAGTGGAGATGATAGACAGCAACTAATGCTG 1864
QY 1493 AAGTTATAAAAACATATTATGCTAATGATCTGAAATAGCTTAAAGCTGCAAGTTTATAGATA 1552
DB 1865 AAGTTTAAAGAGATGATTATTTC----- 1890
QY 1553 AAGATTCAGGCTCTCAAGTAGTAGTCAGGAGTATTTAAATTTCTATGTAGCTAAAGATG 1612
DB 1891 -----ACAGATGGTGAAGTTGTGAATTTACTTTGTTGCAAAAGATG 1930
QY 1613 GATCTACAAAAGAGATCAATTTAGTTGATGCAATTAGCAGTAGCAGCTGTTGCTGG----- 1667
DB 1931 GTTCTACTAAGAGATCAATTTAGTAGATGCCCTTAGCAGCAGCAACCAATAGCAGGTAGAT 1990
QY 1668 ----ATAAATACTGCTCCAGTTGTTATAGCTATGATTTCTTTATCTTCTGATCAATCGG 1723
DB 1991 TTAAGAGAGTCTCCAGCTCCAATCATAGCTAGCTACTGATCTTATCTTCTGACCAAAATG 2050

```



```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16859
; LENGTH: 1893
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-282-122A-16859

Query Match          7.9%; Score 145.2; DB 17; Length 1893;
Best Local Similarity 53.1%; Pred. No. 2.7e-13;
Matches 412; Conservative 0; Mismatches 343; Indels 21; Gaps 4;

QY 863 AAAATGTTGATTAGTAGGTGCAAAATGCACTAGTAGATGGATTAGTTGGCGCTCCTTTAG 922
DB 1037 AAGCTATAATTTTAGTTGGAGAAGATGCTATAGTTGATGGATTAGCATCAGCACCACCTTG 1096

QY 923 CAGCAGAAAAAGATGCTCCACTATTATTAATCTCAAAAAGATTAATAGATTGTCGTCAATGA 982
DB 1097 CCTCTCAAAAAATGACCAATATTATTTATCTAAAAAAGATTCACTACCATCAGAAATAG 1156

QY 983 AATCTGAATAAAGACAGATTTTAGACTTAAAACTTCAACAGAGTAACAGGAAAAACAG 1042
DB 1157 AAGCTGAAATATTAAAGAT-----ACTTGAAGTAACTTATCTCTTAAGAAAA 1204

QY 1043 TTTATATAGCTGGTGAGTTAATAGTGTATCTAAAGAAAGTTGTAAACAGAA---TTAGAAAT 1099
DB 1205 TATATATAGTAGTGAGAACTCTAAAGTATCAAAAGAACTGAAGAAAAAATCTTCTAAT 1264

QY 1100 CAATGGGATTAAGTTGAAAGATTCCTCAGGTGATGATGATGATGAACTCTCTTTAAAAA 1159
DB 1265 TAGGTGTAAGTAAAGTTGACAGAGTTCTCGAAGAGATAGATTGAAACTTCTTTAGAAA 1324

QY 1160 TAGCAGGTGAATAGGCTTAGATAAATGATAAGGCTTATAGTTGGTGAACAGGATTAG 1219
DB 1325 TAGCAAAACAATTAAGAGATACATTTTAAGACTGCTTTTGTAGTAGTGGAATGGAGAAG 1384

QY 1220 CAGATGCCATGATAGTACGTTGCTTCTACTAAATTAGATGGTAAATGGTGTGTAG 1279
DB 1385 CTGATGCTATGAGTATA---TCAGCTAGAGCTGCTCAATTTGGTGTCCCAATATAGTTA 1441
```

```
QY 1280 ATAGAAACAATGGACATGCTACTCAATAGATTGTTGTTAGATGGAAGCTGATAAATAT 1339
DB 1442 CAGGTAACGAATTTAGATGCAAAATGCTGAAAAATTTATTTAAAGGAAAAAGAAATTTAGAAATAG 1501
QY 1340 CTGATGACTTTAGATAGTTTCTTTAGGAAGCGCTGATGTAGATATAATAGTGGATTGCAA 1399
DB 1502 TGGGTG---GAGAAAATTTCTGTATCAAAAGAAGTTGAAGACAAATTTAGTATATAGATT 1558
QY 1400 GTGTATCTGAAAGAATGGAAGAAGCTATATCAGATGCTACTGTTAAAGCGCTTACAAGAG 1459
DB 1559 TAAATAATAAAGTTGAAAGATTTAGCTGGAGAAAAATAGAAAAGATATACTAATGCTAAAGTAA 1618
QY 1460 TTAAGGGCGCATAGACACACTAACTCTGAAGTTTATAAAAACATATTTATGCTAATG 1519
DB 1619 TCAATATAATATCTATCGAGGTGCAACTNAAGCATATGTAGCAAAAGATGGTTATGTAGGTG 1678
QY 1520 ATACTGAAATAGCTAAAGCTGCAGTGTTTAGATAAAGATTTCAAGTGCTTCAAGTAGTGATG 1579
DB 1679 GAAATGGACAATTAGTTGATGCACTTACAGCAGCACCACCTTGACACTAGTTCAAAAGCTC 1738
QY 1580 CAGGAGTATTTAATTTCTTATGTAGCTNAAGATGGATCTTACAAAAGAGATCAATTA 1635
DB 1739 CAATAGTATTAACTACAGAAGAACTTTCTTAAATCAAGAAGAGTAGTTGAGTTA 1794
```

RESULT 13

```
US-10-282-122A-16782
; Sequence 16782, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16782
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Clostridium difficile
```


US-10-282-122A-16782

Query Match 5.78; Score 104; DB 17; Length 1503;
Best Local Similarity 51.8%; Pred. No. 7.3e-07;
Matches 338; Conservative 0; Mismatches 275; Indels 39; Gaps 3;
QY 856 AATGTTAAAAATGTTGTTAGTAGGTGCAATGCACTAGTAGATGATTTAGTTGGCGCT 915
DB 244 AGTGCAATGAGGCTGTGTTGTAATAGTAGTAATCTATTGTAGATGATTTAGTTGTAACA 303
QY 916 CCTTTAGCAGCAGAAAAGATGCTCCACTATTATTAATCTTCAAAAGATATAAATAGATTGCG 975
DB 304 CCATTTGCTAAATTTAAAAAATGCTCTATATTATTGACAGAAAAAATAATCTAAATTTCA 363
QY 976 TCAGTAAATCTGAAATTAAGAGAGTTTGTAGACTTAAAACTTCAACAGAGTAACAGGA 1035
DB 364 CAAACTGAGAAAGAAATTTAAAAGAT-----TAGGAGTT 396
QY 1036 AAAACAGTTTATATAGCTGCTGGAGTTAATAGTGTATCTAAAGAAAGTTGTAACAGAAATTA 1095
DB 397 AAAAAGTTTATATATTTGTTCTATCTAGTTCTGTATCAATGATGTCGCAAAACAGTTA 456
QY 1096 GAATCAATGGGATTTAAAAGTTGAAAGATTTCTCAGGTGATGATAGATGAACTTTCTTTA 1155
DB 457 GAATCTAATTTAACTACTGATAGAGTTGGTGTGTGATAGATACCAACAGCTTTG 516
QY 1156 AAAATAGCAGTGAAATAGGCTTAGATAATGAT-----AAGGCTTATGATTTGTT 1206
DB 517 AGTATAGCTAAAAAATAGAGGATTTAAAAGATATATCAGAAATAGCTGTAGTAATGGG 576
QY 1207 GGAACAGATTAGCAGATGCAATGATAGCTTCACTGCTTCTCTAAATAGATGGT 1266
DB 577 TACACAGCTCTTGACAGATGCGATGATAGCTTCTGTAGCTGCTACTAATGGAATGGT 636
QY 1267 AATGGTGTGTAGATAGAACAAATGGACATGCTACTCCAATAGTTGTTGTAGATGGAATA 1326
DB 637 ATACTTCCAGTATCTGATGATCTGTTATATCTAGCTTTAAAGATTTTATAGAACTTAAG 696
QY 1327 GCTGATAAATATCTGATGATCTTATAGATGTTTCTTAGAGCGCTGATGATAGATATAA 1386
DB 697 AATATAAATTAATCATATAATAGGCTCTACTAATGCTGTATCAGATAAA---ATAAAA 753
QY 1387 GGTGATTTGCAAGTGTATCTGAAAAGATGGAAGATGCAAGATCTATCAGATGCTACTGGTAAA 1446
DB 754 CAAAGTTTACTTAATTCGAAAGAAATAGGTGGACAGATAGAAATGAACTAATGGTAAA 813
QY 1447 GCGGTTACAGAGTTAAAGGCGACGATAGACAGACACTAACTCTGAAGTTA 1498
DB 814 GTAATAGAAAGTTTATACGCTCTAAATAAATTTAAACAATGTATTTGTGGCTA 865

RESULT 14
US-10-282-122A-16845
; Sequence 16845, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16845
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Clostridium difficile
US-10-282-122A-16845
Query Match 5.4%; Score 99.6; DB 17; Length 1830;
Best Local Similarity 53.9%; Pred. No. 3.8e-06;
Matches 275; Conservative 0; Mismatches 199; Indels 36; Gaps 2;
QY 856 AATGTTAAAAATGTTGTTAGTAGGTGCAATGCACTAGTAGATGATTTAGTTGGCGCT 915
DB 154 AATCTGAAAGTCTGCTTATGTAATGACATTCATAGCAGATGTTTATCAGCTACT 213
QY 916 CCTTTAGCAGCAGAAAAGATGCTCCACTATTATTAATCTTCAAAAGATATAAATAGATTGCG 975
DB 214 CCATTTGCTAAACCAAAAGATGCGCAATATTATTAACTCAAAGTATAAATTTGGATAGT 273
QY 976 TCAGTAAATCTGAAATTAAGAGAGTTTGTAGACTTAAAACTTCAACAGAGTAACAGGA 1035
DB 274 AGAACAAAAGCAGAAATTTAAAAGACTTGGTGTG----- 306
QY 1036 AAAACAGTTTATATAGCTGCTGGAGTTAATAGTGTATCTAAAGAAAGTTGTAACAGAAATTA 1095
DB 307 AAAAATGATATTATTAATAGTGGTTCATTTGCATTAAAGTTTCAAGAGATTGAAAGCAATTA 366
QY 1096 GAATCAATGGGATTTAAAAGTTGAAAGATTTCTCAGGTGATGATAGATGAACTTCTTTA 1155
DB 367 AATGCAAAAAATATATAATTTTGAAGAATATCCGGAATAGTAGATATGATATCTTCTTTA 426
QY 1156 AAAATAGCAGTGAAATAGGCTTAGATAATGATAAGGCTTATGTTAGT-----TGCT 1206
DB 427 AAATTAGCTGAAAAGCTAGATAGGGAAGCTCTATATCTAAATAGTAGTAGTAATAGGA 486
QY 1207 GGAACAGGATTAGCAGATGCCATGATAGCTTCACTGCTTCTACTTAAATAGATGGT 1266
DB 487 GAAAGGCTCTTCTGATGATGAGTGGTGGAGCTATAGCTGCTCAAGAAACATGCCA 546
QY 1267 AATGGTGTGTAGATAGAACAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1326
DB 547 ATAATACTTTCTGATTCAGAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 606
QY 1327 GCTGATAAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1356
DB 607 GATATAGCAAAATCGTATGTAATAGGTGGT 636

RESULT 15
US-09-742-096-2
; Sequence 2, Application US/09742096
; Patent No. US20020155441A1

```

; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 5361
; TYPE: DNA
; ORGANISM: P. falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5361)
; OTHER INFORMATION:
US-09-742-096-2

Query Match      5.2%; Score 95.6; DB 9; Length 5361;
Best Local Similarity 43.6%; Pred. No. 2.2e-05;
Matches 645; Conservative 0; Mismatches 809; Indels 24; Gaps 4;

QY      233 CAATGAGAAAGATTCTTCAACTCCAACTGGAGATAAAGTTTATTTCAGAACAACTT 292
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      293 TAACTACAGCTAATGAAATGAAGATTATGTAAGAACA---CTTTAAAAATTTAGATG 349
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      350 CAGGAGAAATGCTATATAGATTAACTTATATATATATGCTAAACTGTTGAAATTAAG 409
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      890 CAAGTGTGGAAGAAAGGTGAGAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTTGAAG 949
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      410 TAGTAGCAGCTAGTGAAGAAACAGTAGTGTATCTAGTATGCGAAATATAGTGCAAAAG 469
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      950 AAAGTGTAGCTGAAATGTTGAAGAAAGGTAGCTGAAATGTTGAAGAAAGTGTAGCTG 1009
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      470 ATATAGCTGAAATATGTTTGAAGACAAAGACTTGAAGAAATGCACTAAAAACTATAA 529
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1010 AAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1069
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      530 ATGCTTCAGATTTCAGTAAACTGATGATGATGATGATGATGATGATGATGATGATGATG 589
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1070 AAAATGTAGCTCCAAAGTGTGTAGAAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAAG 1129
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      590 AGAGATTACAGGTTTCTCAACTTATAGAGCTACAAATATATATGAGGAACTGCGATAG 649
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1130 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAG 1189
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      650 GTAATACACAGTAATATTAACCTCTAAATCTACTAGTAAGAGTAAATTTAAAGACTGCGAG 709
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1190 AAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTG 1249
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      710 TAGAAGAGTTACAAAAATGCAATGCTAGTTATTTCTAATACTACAACTTTAGCTGGTGAAG 769
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1250 AAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1309
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      770 ACAGAAATACAAACAGCTATAGAGATAAGTAAAGAAATATTACAAATATGATGGCGAGAAAT 829
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1310 AAATGTAGCTCCAAAGTGTGTAGAAGTGTGCTCCAAAGTGTGTAGAAGTGTAGAAG 1369
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      830 CAGATCATTCAGCTGATGTTTAAAGAGAAATGTTAAAAATGTTTATTTAGTAGGTGCAAAATG 889
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1370 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTTGAAG 1429
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY      890 CACTAGTAGATGGAATTAGTTGGGCTCCTTTAGCAGCAGCAAAAAGATGCTCCACTATTAT 949
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1430 AAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTG 1489
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      950 TAACTTCAAAAGATAAATTAGATTTCGTGAGTAA-----ATCTGAAATAAAGAGAG 1000
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1490 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAG 1549
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1001 TTTTAGACTTAAAACTTCAACAGAAAGTAAACAGGAAACAGATTATATAGCTGTGGAG 1060
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1550 AAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTTAGCTC 1609
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1061 TTAATAGTGTATCTAAAGAGTGTAAACAGAAATAGAAATCAATGGGATTAAAGTTGAAA 1120
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1610 CAAAGTGTGTAGAAAGTGTGGCTCCAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAG 1669
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1121 GATTCTCAGGTGATGATAGATATAAATCTTTTAAAAATAGCAGGTGAAATPAGCTTAG 1180
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1670 AAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTG 1729
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1181 ATAATGATAAGGCTTTATGTTGGTGAACAGGATTAAGCATGCCATGAGTATAGCTT 1240
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1730 AAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAG 1789
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1241 CAGTTGCTTCTACTAAATTAGATGCTAATG-----GTGTTGTAGATAGAACAAATG 1291
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1790 AAATCGTAGCTCCAACTGTTGAAGAAATTTAGCTCCAAAGTGTGTAGAAAGTGTGGCTC 1849
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1292 GACATGCTACTCCAATAGTTGTTAGATGGAAGAAAGTGTAAAAATATCTGATGACTTAG 1351
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1850 CAAAGTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAG 1909
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1352 ATAGTTTCTTAGGAAGCGCTGATGTAGATATATATAGGTGGATTTCGAAGTGTATCTGAAA 1411
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1910 AAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTC 1969
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1412 AGATGGAAGAGCTATATCAGATGCTACTGTGTAAGGCGTTTCAAGAGTTTAAAGCGACG 1471
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1970 CAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTTAGCTCCAAGTGTGTAG 2029
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1472 ATAGCAACAGACACTCAACTCTGAAGTTTATAAAACAATTTATGCTAATGATCTGAAATAG 1531
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      2030 AAAGTGTGGCTCCAAGTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAG 2086
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1532 CTAAAGCTGCGAGTTTATAGATAAGATTCAGGTGCTTCAAGTGTGATCGCAGAGTATTTA 1591
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      2087 CTGAAAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTG 2146
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1592 ATTTCTATGTAGCTAAAGTGGATCTACAAAGAGAGATCAATTAGTTGATGATTAGCAG 1651
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      2147 AAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAG 2206
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      1652 TAGGAGCTGTTCTCGATATATAAATTTGCTCCAGTTGTA 1689
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      2207 CTCCAACTGTTGAAGAAATTTAGCTCCAAAGTGTGTA 2244
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: October 6, 2005, 11:30:56
Job time : 3714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM nucleic - nucleic search, using sw model
Run on: October 6, 2005, 05:21:23 ; Search time 7980 Seconds
(without alignments)
11111.910 Million cell updates/sec

Title: US-10-068-870A-4
Perfect score: 1830
Sequence: 1 agaaaaaagaatttagc.....tgaaagatttagatagatg 1830

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 4708233 seqs, 24227607955 residues 9416466
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	1830	100.0	AX512902 Sequence
2	1815.6	99.2	AP478570 Clostridi
3	1159.6	63.4	AP458877 Clostridi
4	1096.2	59.9	AP458878 Clostridi
5	546.6	29.9	AY004256 Clostridi
6	542	29.6	AX512904 Sequence
7	535.6	29.3	AJ300676 Clostridi
8	535.6	29.3	AX259458 Sequence
9	503.4	27.5	AX512906 Sequence
10	483.2	26.4	CDI300677 Clostridi
11	483.2	26.4	AX259456 Sequence
12	480	26.2	AX512907 Sequence
13	471.8	25.8	AX512901 Sequence
14	471.8	25.8	AX512903 Sequence
15	471.8	25.8	AX512908 Sequence
16	471.8	25.8	AJ291709 Clostridi
17	471.8	25.8	AX259457 Sequence
18	470.2	25.7	AX512905 Sequence
19	453.2	24.8	AP478571 Clostridi

20	135	7.4	253001	3	AE014834
21	119.4	6.5	208332	10	AC122227
22	117.4	6.4	5420	3	AF056936
23	112.4	6.1	2069	6	E10125
24	112.4	6.1	3399	6	E10126
25	112.4	6.1	347050	3	PF929351
26	112.2	6.1	16377	3	AF202180
27	112	6.1	4995	3	PFAMESA
28	111.8	6.1	271546	3	AE014843
29	110.6	6.0	293650	1	AP004172
30	109.4	6.0	237549	2	AC112451
31	108	5.9	4102	3	AF270648
32	107.8	5.9	170221	10	AC127421
33	107.4	5.9	233713	2	AC133034
34	104.4	5.7	110000	2	AC129389
35	103.4	5.7	172307	2	AC044842
36	103.2	5.6	202920	10	AL929441
37	103	5.6	219359	10	AC116998
38	101.6	5.6	179685	10	AC122056
39	101.2	5.5	4507	1	AF093106
40	100	5.5	34347	2	BX957360
41	99.6	5.4	1833	1	AF194870
42	98.6	5.4	169614	5	BX511115
43	98.6	5.4	177122	2	BX57314
44	96.8	5.3	165529	10	AC147621
45	96	5.2	219181	2	AC141063

ALIGNMENTS

RESULT 1
AX512902
LOCUS AX512902 1830 bp DNA linear PAT 03-OCT-2002
DEFINITION Sequence 2 from Patent WO02062379.
ACCESSION AX512902
VERSION AX512902.1 GI:23504061
KEYWORDS Clostridium difficile
SOURCE Clostridium difficile
ORGANISM Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE Doyle, R., Kelleher, D., Windle, H.J., Walsh, J.B. and Deirdre, N.B.
AUTHORS
TITLE Clostridium difficile vaccine
JOURNAL Patent, 02062379-A, 2 15-AUG-2002;
THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH (IE)

FEATURES
Location/Qualifiers
source 1. .1830
/organism="Clostridium difficile"
/mol_type="unassigned DNA"
/db_xref="taxon:1496"
ORIGIN
Query Match 100.0%; Score 1830; DB 6; Length 1830;
Best Local Similarity 100.0%; Pred. No. 4e-240;
Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATCAAAAAGAAATTTAGCAATGGCTGATGGCAGCTGTTACTGTAGTAGTCTGCTGCT 60
1 ATCAAAAAGAAATTTAGCAATGGCTGATGGCAGCTGTTACTGTAGTAGTCTGCTGCT 60
Db 1 ATCAAAAAGAAATTTAGCAATGGCTGATGGCAGCTGTTACTGTAGTAGTCTGCTGCT 60
Qy 61 CCAGTTTTCACAGCTTCAGATGTAATATCACTACAGATGTCACAAATGATAAGTAT 120
61 CCAGTTTTCACAGCTTCAGATGTAATATCACTACAGATGTCACAAATGATAAGTAT 120
Db 61 CCAGTTTTCACAGCTTCAGATGTAATATCACTACAGATGTCACAAATGATAAGTAT 120
Qy 121 ACAGTATCAATTAAGTACTAGTACTAGTAAAGGATATTTTAGCAGCACAACCTTA 180
121 ACAGTATCAATTAAGTACTAGTACTAGTAAAGGATATTTTAGCAGCACAACCTTA 180
Db 121 ACAGTATCAATTAAGTACTAGTACTAGTAAAGGATATTTTAGCAGCACAACCTTA 180
Qy 181 ACAACAGGTCAGTATTTTGAACAAAGATACAAAAGTTACTTTCTATGATGCAATGAG 240
181 ACAACAGGTCAGTATTTTGAACAAAGATACAAAAGTTACTTTCTATGATGCAATGAG 240

Db 181 ACAACAGTGCAGTTATTTTGAACAAAGATACAAAGTTACTTTCTATGATGCAATGAG 240
QY 241 AAAGATTCCTCAACTCCAACTGGAGATAAAAAAGTTTATTCAGAACTTTTAACTACA 300
Db 241 AAAGATTCCTCAACTCCAACTGGAGATAAAAAAGTTTATTCAGAACTTTTAACTACA 300
QY 301 GCTAATGGAATCGAGATTATGTAAGACAACTTTTAAAAAATTTAGATGCGAGAGATAT 360
Db 301 GCTAATGGAATCGAGATTATGTAAGACAACTTTTAAAAAATTTAGATGCGAGAGATAT 360
QY 361 GCTATTATAGATTAACTTTATTAATATGCTTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT 420
Db 361 GCTATTATAGATTAACTTTATTAATATGCTTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT 420
QY 421 AGTGAACAAACAGTAGTTCTTCTAGTGATGCGAAAAATAGTGCAAAAAGATATAGCTGAA 480
Db 421 AGTGAACAAACAGTAGTTCTTCTAGTGATGCGAAAAATAGTGCAAAAAGATATAGCTGAA 480
QY 481 AAATATGCTGTTGAAGACAAAGACTTAGAATATGCACTAAAAAATCTAATAATGCCTCAGAT 540
Db 481 AAATATGCTGTTGAAGACAAAGACTTAGAATATGCACTAAAAAATCTAATAATGCCTCAGAT 540
QY 541 TTCAAGTAAACCTGATAGTTACTATCAAGTAGTTCTTTATCCAAAAAGGAAAGATTAACA 600
Db 541 TTCAAGTAAACCTGATAGTTACTATCAAGTAGTTCTTTATCCAAAAAGGAAAGATTAACA 600
QY 601 GGTTCCTCAACTTATAGAGCTCAAAATTAATGAAGGAACTGCATATAGTGAATACACA 660
Db 601 GGTTCCTCAACTTATAGAGCTCAAAATTAATGAAGGAACTGCATATAGTGAATACACA 660
QY 661 GTAATATTAACCTTAAATCTACTAGTAGAGTTAATTTAAAGCTGCAGTAGAAGAGTTA 720
Db 661 GTAATATTAACCTTAAATCTACTAGTAGAGTTAATTTAAAGCTGCAGTAGAAGAGTTA 720
QY 721 CAAAAATGAATGCTAGTTATTTCTAATACTACAACTTTAGCTGGTGATGACAGAAATACA 780
Db 721 CAAAAATGAATGCTAGTTATTTCTAATACTACAACTTTAGCTGGTGATGACAGAAATACA 780
QY 781 ACAGCTATAGAGTAAGTAAAGATTAATTAATATGATGGCGAGAAATCAGATCAATCA 840
Db 781 ACAGCTATAGAGTAAGTAAAGATTAATTAATATGATGGCGAGAAATCAGATCAATCA 840
QY 841 GCTGATGTTAAAGAGAACTTAAATGTTGTTATAGTAGTGCAAAATGCACCTAGTAGAT 900
Db 841 GCTGATGTTAAAGAGAACTTAAATGTTGTTATAGTAGTGCAAAATGCACCTAGTAGAT 900
QY 901 GGATTAGTTGCGGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTCAAAA 960
Db 901 GGATTAGTTGCGGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTCAAAA 960
QY 961 GATAAATTAGATTCGTCACTAAATCTGAAATAAAGAGAGTTTATAGACTTAAAACTTCA 1020
Db 961 GATAAATTAGATTCGTCACTAAATCTGAAATAAAGAGAGTTTATAGACTTAAAACTTCA 1020
QY 1021 ACAGAGTAAACAGGAAAAACAGTTATATAGCTGGTGAGTTAATAGTGTATCTAAAGAA 1080
Db 1021 ACAGAGTAAACAGGAAAAACAGTTATATAGCTGGTGAGTTAATAGTGTATCTAAAGAA 1080
QY 1081 GTTGTAAACAGATTAGAATCAATGGGATTTAAAGTTGAAAGATTTCTCAGGTGATGAGA 1140
Db 1081 GTTGTAAACAGATTAGAATCAATGGGATTTAAAGTTGAAAGATTTCTCAGGTGATGAGA 1140
QY 1141 TATGAAACCTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGGCTTATGTA 1200
Db 1141 TATGAAACCTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGGCTTATGTA 1200
QY 1201 GTTGTGGAACAGGATTAGCAGATGCCATGAGTAGCTTCAAGTTGCTTCTACTAAATTA 1260
Db 1201 GTTGTGGAACAGGATTAGCAGATGCCATGAGTAGCTTCAAGTTGCTTCTACTAAATTA 1260
QY 1261 GATGGTAATGGTGTGTAGATAGAACAAATGGACATGCTACTCCCAATAGTTGTTGTAGAT 1320
Db 1261 GATGGTAATGGTGTGTAGATAGAACAAATGGACATGCTACTCCCAATAGTTGTTGTAGAT 1320

QY 1321 GGAAAGCTGATATAAATATCTGATGACTTAGATAGTTTCTTAGAGAGCGCTGATGTAGAT 1380
Db 1321 GGAAAGCTGATATAAATATCTGATGACTTAGATAGTTTCTTAGAGAGCGCTGATGTAGAT 1380
QY 1381 ATAATAGCTGATTTGCAAGTGATCTCGAAAGATGGAAGAGCTATATCAGATGCTACT 1440
Db 1381 ATAATAGCTGATTTGCAAGTGATCTCGAAAGATGGAAGAGCTATATCAGATGCTACT 1440
QY 1441 GGTAAAGCGTTTACAAGGTTTAAAGCGCAGATAGACAAAGACACTAACTCTGAAGTTATA 1500
Db 1441 GGTAAAGCGTTTACAAGGTTTAAAGCGCAGATAGACAAAGACACTAACTCTGAAGTTATA 1500
QY 1501 AAAACATATTTATGCTAATGATACCTGAATAGCTAAAGCTGCAGTTTATAGATAAGATTCA 1560
Db 1501 AAAACATATTTATGCTAATGATACCTGAATAGCTAAAGCTGCAGTTTATAGATAAGATTCA 1560
QY 1561 GGTGCTTCAAGTAGTAGTGATGAGGAGTATTTAATTTCTATGTAGCTAAAGATGGATCTACA 1620
Db 1561 GGTGCTTCAAGTAGTAGTGATGAGGAGTATTTAATTTCTATGTAGCTAAAGATGGATCTACA 1620
QY 1621 AAAGAAGATCAATTAGTTGATGCACTTAGCAGTAGGAGCTGTTGCTGGATATAAACTTCT 1680
Db 1621 AAAGAAGATCAATTAGTTGATGCACTTAGCAGTAGGAGCTGTTGCTGGATATAAACTTCT 1680
QY 1681 CCAGTTGTTAGCTACTGATTTCTTATCTTCTGATCAATCGTTGCTTATTAAGCAAGTT 1740
Db 1681 CCAGTTGTTAGCTACTGATTTCTTATCTTCTGATCAATCGTTGCTTATTAAGCAAGTT 1740
QY 1741 GTAGGAGAAAAATATTTCAAAGATTTAAACAAAGTTTGAAGAGTCAAGTAATAGCTAAATTCAGTT 1800
Db 1741 GTAGGAGAAAAATATTTCAAAGATTTAAACAAAGTTTGAAGAGTCAAGTAATAGCTAAATTCAGTT 1800
QY 1801 ATAAACAAATGAAAGATTTTATTAGATG 1830
Db 1801 ATAAACAAATGAAAGATTTTATTAGATG 1830
RESULT 2
AF478570
LOCUS AF478570 1833 bp DNA linear BCT 27-JUN-2002
DEFINITION Clostridium difficile surface layer protein A precursor, gene, complete cds.
ACCESSION AF478570
VERSION AF478570.1 GI:21311577
KEYWORDS
SOURCE
ORGANISM Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE 1 (bases 1 to 1833)
AUTHORS Calabi,E. and Fairweather,N.
TITLE Patterns of sequence conservation in the S-Layer proteins and related sequences in Clostridium difficile
J. Bacteriol. 184 (14), 3886-3897 (2002)
JOURNAL 22077258
MEDLINE 12081960
REFERENCE 2 (bases 1 to 1833)
AUTHORS Calabi,E. and Fairweather,N.F.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2002) Biological Sciences, Imperial College, Exhibition Rd, London, Greater London SW7 2AZ, UK
FEATURES
source
1. .1833
/organism="Clostridium difficile"
/mol_type="genomic DNA"
/strain="167"
/db_xref="taxon:1496"
1. .1833
/note="SlpA precursor"
/codon_start=1
/transl_table=11
/product="surface layer protein A precursor"
CDS

/protein id="AAM46789.1"
/db_xref="GI:21311578"
/translation="MKRLNLAAMAAYVVGSAAPVFAAASDVLSIQGTNDKTYVSN
TKASDLVKDILAAQNLITGAVILNKOTKTPFDANEKSSSTFGKKYSEQLITAN
GNEDVYKTLKNDAGEYAIIDLTNNAKTVBIKVAASEKTVVSSDAKNSAEIDAE
KTVFEDKJLENALKTINADSPFKTDSYQVLYPKRQLQFSTRYRANNEGTAYEN
TPVILTKSTKSNLKTAVEELQKLNASYNTTLAGDDRIQTAEISKEYNNDEK
QDHSADVKNVNLVGNALVDGLVAAPLAAEKDAPLLLTSKDLDSVSKSEIKRV
LDLKTSTVTKTVLGGVNSVSKVETELSMGLKVERFSGDRYETSLKIADEIG
LNDKAVVGGTGLADAMSIAVSASTKLDSNGVVDKTHGHPPIVVDGKDKISDDL
DSFLGADYDI11GGFASVSEKMEELSDATGKGVTRKNGDQDNTSEVIKTYIANDT
ETAKAALVDGSSDAGVFNFYAKDGSQKBDQLVDALAVGAVAGYKLPVVLAT
DSLSSDQSVIAISKVGEKYSKDLTVQVQGIANSINKIKDLIDM"

ORIGIN

Query Match		99.2%; Score 1815.6; DB 1; Length 1833;
Best Local Similarity		99.5%; Pred. No. 3.7e-238;
Matches 1821; Conservative 0; Mismatches 9; Indels 0; Gaps 0;		
QY	1	ATCAAAAAAGAAATTTAGCAATGGCTATGCGCAGCTGTTACTGTAGTAGGTTCTGCTGCT 60
DB	1	ATCAAAAAAGAAATTTAGCAATGGCTATGCGCAGCTGTTACTGTAGTAGGTTCTGCTGCT 60
QY	61	CCAGTCTTTTGCAGCAGCTCCAGATGTAATATCACTACAAGATGGTACAAATGATAAGTAT 120
DB	61	CCAGTCTTTTGCAGCAGCTCCAGATGTAATATCACTACAAGATGGTACAAATGATAAGTAT 120
QY	121	ACAGTATCAAAATCTAAAGCTAGTGAATAGTAAAGGATATTTTACGAGCACAACACTTA 180
DB	121	ACAGTATCAAAATCTAAAGCTAGTGAATAGTAAAGGATATTTTACGAGCACAACACTTA 180
QY	181	ACAACAGGTCAGTATTTTGAACAAGATACAAAGTTACTTTCTATGATGCAAAATGAG 240
DB	181	ACAACAGGTCAGTATTTTGAACAAGATACAAAGTTACTTTCTATGATGCAAAATGAG 240
QY	241	AAAGATCTTCAACTCCAGCTGAGATATAAAGTTTATTCAGAACAACTTTTAACTACA 300
DB	241	AAAGATCTTCAACTCCAGCTGAGATATAAAGTTTATTCAGAACAACTTTTAACTACA 300
QY	301	GCTAATGGAAATGAAGATATGTAAGACAACTTTTAAAGATTTAGATGCGAGAGATAT 360
DB	301	GCCAAATGGAATGAAGATATGTAAGACAACTTTTAAAGATTTAGATGCGAGAGATAT 360
QY	361	GCTATTATAGATTTAACTATAATATGCTAAAGCTGTTGCAAAATTAAGTAGTAGAGCT 420
DB	361	GCTATTATAGATTTAACTATAATATGCTAAAGCTGTTGCAAAATTAAGTAGTAGAGCT 420
QY	421	AGTGAAAAAAGCAGTAGTGTATCTAGTGTGCGAAAAATAGTGCAAAAATATAGCTGAA 480
DB	421	AGTGAAAAAAGCAGTAGTGTATCTAGTGTGCGAAAAATAGTGCAAAAATATAGCTGAA 480
QY	481	AAATATGTTTGAAGACAAAGACTTAGAAAAATGCACTAAAACTATAAATGCTCAGAT 540
DB	481	AAATATGTTTGAAGACAAAGACTTAGAAAAATGCACTAAAACTATAAATGCTCAGAT 540
QY	541	TTCACTAAAACTGATAGTACTATCAAGTAGTCTTTTATCCAAAAGGAAAGAGATTACAA 600
DB	541	TTCACTAAAACTGATAGTACTATCAAGTAGTCTTTTATCCAAAAGGAAAGAGATTACAA 600
QY	601	GGTTTCTCAACTATAGAGCTACAAATTAATAAGAGAACTGCATATGCTAATACACCA 660
DB	601	GGTTTCTCAACTATAGAGCTACAAATTAATAAGAGAACTGCATATGCTAATACACCA 660
QY	661	GTAATATTAACTCTAAATCTACTAGTAAGAGTAAATTTAAAGCTGCAAGTAGAGGTTA 720
DB	661	GTAATATTAACTCTAAATCTACTAGTAAGAGTAAATTTAAAGCTGCAAGTAGAGGTTA 720
QY	721	CAAAAATTTGAATGCTAGTATTCTAATCTACAACTTTAGCTGTGATGACAGATACAA 780
DB	721	CAAAAATTTGAATGCTAGTATTCTAATCTACAACTTTAGCTGTGATGACAGATACAA 780
QY	781	ACAGCTATAGAGATAAGTAAAGATATTTACAAATAGTGCAGAGAAATCAGATCATTTCA 840

RESULT 3
AP458877

LOCUS AF458877 1182 bp DNA linear BCT 07-JUL-2002
DEFINITION Clostridium difficile strain ATCC 43594 S-layer protein variable domain SlpA (slpA) gene, partial cds.
ACCESSION AF458877
VERSION AF458877.1 GI:21702530
KEYWORDS
SOURCE
ORGANISM Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE 1 (bases 1 to 1182)
AUTHORS Kartjalainen, T., Saumier, N., Barc, M.C., Delmee, M. and Collignon, A.
TITLE Clostridium difficile Genotyping Based on slpA Variable Region in S-Layer Gene Sequence: an Alternative to Serotyping
JOURNAL J. Clin. Microbiol. 40 (7), 2452-2458 (2002)
MEDLINE 22083941
PUBMED 12089261
REFERENCE 2 (bases 1 to 1182)
AUTHORS Kartjalainen, T.K. and Saumier, N.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-2001) Microbiology, University of Paris-Sud, Faculty of Pharmacy, 5, rue JB Clement, Chateauf-Malabry 92296, France

FEATURES
Location/Qualifiers
1..1182
/organism="Clostridium difficile"
/mol_type="genomic DNA"
/strain="ATCC 43594"
/db_xref="ATCC:43594"
/db_xref="taxon:1496"
/notice="serogroup: A1"
<1..>1182
/gene="slpA"
/gene="slpA"
/codon_start=1
/transl_table=11
/product="S-layer protein variable domain SlpA"
/protein_id="AAW75941.1"
/db_xref="GI:21702531"
/translation="AAPVFAASDVISLQDGTNRYTVSNTKASDLVKDILAAQNLTT
GAIVLKDQTVFYDANEDKDSPTGDKVYSEOTLTANGNEDVYVTKLNLDA
AIIIDTNNAKTVEIKVVAASEKTVVSSDAKNSAKIAEKVYFEDKDLKWLPLKIN
ASDFKTSYQVVLVYPKGRLOGFSTVATYNGTAYGNTPVILILKSTSKNLKT
AVELOKLNVSNTITLAGDRIQTAIEISKBYTNDGERSDHSADKVENKNVLY
GANLVLDGLAAPAAEKDAPLLTSKDLSSVKSEIKRLDLKSTSTVTKTVYIA
GGVNSVAKVVEVTELESMLKVERFSGDDRYETSLKIADEIGLDNDKAYVVGGTGLADA
MS"

ORIGIN

Query Match 63.4%; Score 1159.6; DB 1; Length 1182;
Best Local Similarity 99.4%; Pred. No. 9.1e-149;
Matches 1175; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
QY 55 GCTGCTCCAGTTTTCGACGAGCTTCAGATGAATATCACTACAGATGGTACAAATGAT 114
DB 1 GCTGCTCCAGTTTTCGACGAGCTTCAGATGAATATCACTACAGATGGTACAAATGAT 60
QY 115 AAGTATACAGTATCAAACTAAAGCTAGTGAATAGTAAAGGATATTTTAGCAGACAA 174
DB 61 AAGTATACAGTATCAAACTAAAGCTAGTGAATAGTAAAGGATATTTTAGCAGACAA 120
QY 175 AACTTAAACACAGGTGCAGTTATTTTGAACAAAGATACAAAGTTACTTTCTATGATGA 234
DB 121 AACTTAAACACAGGTGCAGTTATTTTGAACAAAGATACAAAGTTACTTTCTATGATGA 180
QY 235 AATGAGAAAGATCTTCAACTCACTCGAGATATAAGAGTTTATTCAGAACAACTTTA 294
DB 181 AATGAGAAAGATCTTCAACTCACTCGAGATATAAGAGTTTATTCAGAACAACTTTA 240
QY 295 ACTACAGCTAAATGGAATGAAGATATATGTAAGACAACTTTTAAAAAATTTAGATGACGA 354
DB 241 ACTACAGCTAAATGGAATGAAGATATATGTAAGACAACTTTTAAAAAATTTAGATGACGA 300

QY 355 GAATATGCTATTATAGATTAACTTATTAATATGCTAAAACTGTTGAAATTAAGTAGTA 414
DB 301 GAATATGCTATTATAGATTAACTTATTAATATGCTAAAACTGTTGAAATTAAGTAGTA 360
QY 415 GCAGCTAGTGAAGAAAAACAGTAGTGTATCTAGTGTATGCGAAAAATAGTGCAGAAAGATATA 474
DB 361 GCAGCTAGTGAAGAAAAACAGTAGTGTATCTAGTGTATGCGAAAAATAGTGCAGAAAGATATA 420
QY 475 GCTGAAAAATATGTTGTTTGAAGACAAAGACTTTAGAAAA---TGCACTAAAAAATATATAAT 531
DB 421 GCTGAAAAATATGTTGTTTGAAGACAAAGACTTTAGAAAAATGCGCACTAAAAAATATATAAT 480
QY 532 GCCTCAGATTTTCAGTAAAAACGATGTTACTATCAAGTAGTTCCTTTATCCAAAAAGGAAG 591
DB 481 GCCTCAGATTTTCAGTAAAAACGATGTTACTATCAAGTAGTTCCTTTATCCAAAAAGGAAG 540
QY 592 AGATTACAGGTTTCTCAAACTTATAGAGCTACAAATATTAATGAAGAACTGCATATGCT 651
DB 541 AGATTACAGGTTTCTCAAACTTATAGAGCTACAAATATTAATGAAGAACTGCATATGCT 600
QY 652 AATACACAGTAAATTAATTAACCTCTAAAACTACTAGTAAGATTAATTTAAAGACTGCAGTA 711
DB 601 AATACACAGTAAATTAATTAACCTCTAAAACTACTAGTAAGATTAATTTAAAGACTGCAGTA 660
QY 712 GAAGAGTTACAAAAATTCGAATGCTAGTTTATTTCTAATACTACAACTTTAGCTGGTGCATGAC 771
DB 661 GAAGAGTTACAAAAATTCGAATGCTAGTTTATTTCTAATACTACAACTTTAGCTGGTGCATGAC 720
QY 772 AGAATACAAAACAGCTATAGAGATAAGTAAAGAAATATTAACAATATATGATGCGAGAGAAATCA 831
DB 721 AGAATACAAAACAGCTATAGAGATAAGTAAAGAAATATTAACAATATATGATGCGAGAGAAATCA 780
QY 832 GATCAATTCAGCTGATGTTTAAAGAGAAATGTTTAAAAATGTTTATTAATAGTGGTGCAGAAATGCA 891
DB 781 GATCAATTCAGCTGATGTTTAAAGAGAAATGTTTAAAAATGTTTATTAATAGTGGTGCAGAAATGCA 840
QY 892 CTAGTAGTGAATGATGTTGCGCTCTTTAGCAGCAGAGAAAGATGCTCCCACTATTATTA 951
DB 841 CTAGTAGTGAATGATGTTGCGCTCTTTAGCAGCAGAGAAAGATGCTCCCACTATTATTA 900
QY 952 ACTTCAAAAGATAAATAGATTTCGTGCTAGTAAAAATCTGAAATATAAGAGAGTGTTTAGACTTTA 1011
DB 901 ACTTCAAAAGATAAATAGATTTCGTGCTAGTAAAAATCTGAAATATAAGAGAGTGTTTAGACTTTA 960
QY 1012 AAAACTTCAACAGAGTAAACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGA 1071
DB 961 AAAACTTCAACAGAGTAAACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGA 1020
QY 1072 TCTAAGAGAGTTGTAACAGAAATTAAGATCAATGGATTTAAAGTTTGAAGATTTCTCAGGT 1131
DB 1021 GCTAAGAGAGTTGTAACAGAAATTAAGATCAATGGATTTAAAGTTTGAAGATTTCTCAGGT 1080
QY 1132 GATGATAGATGATAAACTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAAATGATAAG 1191
DB 1081 GATGATAGATGATAAACTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAAATGATAAG 1140
QY 1192 GCTTATGATGTTGGTGGAAACAGGATTAGCAGATGCCATGAGT 1233
DB 1141 GCTTATGATGTTGGTGGAAACAGGATTAGCAGATGCCATGAGT 1182
RESULT 4
AF458878
LOCUS
DEFINITION Clostridium difficile strain TO005 S-layer protein variable domain SlpA (slpA) gene, partial cds.
ACCESSION AF458878
VERSION AF458878.1 GI:21702532
KEYWORDS
SOURCE
ORGANISM Clostridium difficile
Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

Clostridium.
1 (bases 1 to 1185)
Karjalainen, T., Saumier, N., Barc, M.C., Delmee, M. and Collignon, A.
Clostridium difficile Genotyping Based on *slpA* Variable Region in
S-Layer Gene Sequence: an Alternative to Serotyping
J. Clin. Microbiol. 40 (7), 2452-2458 (2002)
22083941
MEDLINE
PUBMED
12089261
2 (bases 1 to 1185)
Karjalainen, T.K. and Saumier, N.
Direct Submission
Submitted (16-DEC-2001) Microbiology, University of Paris-Sud,
Faculty of Pharmacy, 5, rue JB Clement, Chatenay-Malabry 92296,
France
FEATURES
source
Location/Qualifiers
1. .1185
/organism="Clostridium difficile"
/mol_type="genomic DNA"
/strain="TO005"
/db_xref="taxon:1496"
/note="serogroup: A10"
<1. .>1185
/gene="slpA"
<1. .>1185
/gene="slpA"
/codon_start=1
/transl_table=11
/product="S-layer protein variable domain *SlpA*"
/protein_id="AA075942.1"
/db_xref="GI:21702533"
Translations: AAFPVFAAASDVLSQDGTNDKVTNSKASDLVDKDIILAAQNIIT
GALCRKQTSYFLCHQMKWSSTATLEIKVYSEGLITTTANGNEDYVKTKLNLDA
EYALIDLTNNAKTVELKVVAASEKTVVSSDAKNSAKDIAEKYVPEKDLLENALAKTI
NASDFSKTDSYIQVLLPKGRLOGFSRTATNNEGTAICYPVILTLKTSKSNLK
TAYVEELKLNASNTTTLAGDDRIQTALBISKEYYNNDEKQDHSADVKNVNVV
VGNALVDGIVAAPLAAEKDAPLLLSKDLDSVKSEIKRVLDLTKTSTEVTKTVYI
AGGVNSVSKVWTELESMLKVERFSGDDRYETSLKIADIEGLDNKAYVVGTTGLAD
AMS"
ORIGIN
Query Match 59.9%; Score 1096.2; DB 1; Length 1185;
Best Local Similarity 97.6%; Pred. No. 3.9e-140;
Matches 1156; Conservative 0; Mismatches 23; Indels 6; Gaps 4;
Qy 55 GCTGCTCAGTTTTCGACGAGCTTCAGATGTAATATACATACAGATGCTACAAATGAT 114
Db 1 GCTGCTCAGTTTTCGACGAGCTTCAGATGTAATATACATACAGATGCTACAAATGAT 60
Qy 115 AAGTATACAGTATCAATATCTAAAGCTAGTGTAGTAAAGGATATTTTAGCAGACAA 174
Db 61 AAGTATACAGTATCAATATCTAAAGCTAGTGTAGTAAAGGATATTTTAGCAGACAA 120
Qy 175 AACTTAACAACAGCTGCGATTTATTTTGAAACAAGATACA-AAAGTTACTTTCTATG-ATG 232
Db 121 AACTTAACAACAGCTGCGATTTGTTAGAACAAAGGATACAGAAAGTTACTTTCTATGATG 180
Qy 233 CAATGAGAAAGAT-TCTTCACTCCAACTGCGA---TAAAGAGTTTATTTCAGACAA 288
Db 181 CAATGAGAAAGATGCTTCACTGCGCACTGCGCACTGGAAGTTTATTTCAGACAA 240
Qy 289 ACTTTACTACAGCTATGGAAGTAAAGATATGTAAGACACACTTTTAAAGAAATTTAGAT 348
Db 241 ATTTTAACTACAGCTATGGAAGTAAAGATATGTAAGACACACTTTTAAAGAAATTTAGAT 300
Qy 349 GCAGAGAAATATGCTATTATAGATTTTAACTTATTAATATGTAAGAACTTTGAAATTTAA 408
Db 301 GCAGAGAAATATGCTATTATAGATTTTAACTTATTAATATGTAAGAACTTTGAAATTTAA 360
Qy 409 GTAGTAGCAGCTAGTGAAGAAACAGTAGTTGATCTAGTGTGCGAAATAGTGCAAAA 468
Db 361 GTAGTAGCAGCTAGTGAAGAAACAGTAGTTGATCTAGTGTGCGAAATAGTGCAAAA 420
Qy 469 GATATAGCTGAAAATATGCTGTTTGAAGACAAAGACTTTAGAAATGCACTAAAAACTATA 528

Db 421 GATATAGCTGAAAATATATGTTTGAAGACAAAGACTTTAGAAAATGCACTAAAACTATA 480
Qy 529 AATGCTCAGATTTTCAGTAAAACTGATAGTTACTACTCAAGTAGTTCTTTTATCCAAAGGA 588
Db 481 AATGCTCAGATTTTCAGTAAAACTGATAGTTACTACTCAAGTAGTTCTTTTATCCAAAGGA 540
Qy 589 AAGAGATTACAAGGTTTCTCAACTATATAGAGCTACAAATATTAATGAAGAACTGCATAT 648
Db 541 AAGAGATTACAAGGTTTCTCAACTATATAGAGCTACAAATATTAATGAAGAACTGCATAT 600
Qy 649 GGTAAATACACCACTAATTAATTAATCTTAATCTAGTAAAGTAATTTAAAGACTGCA 708
Db 601 GGTAAATACACCACTAATTAATTAATCTTAATCTAGTAAAGTAATTTAAAGACTGCA 660
Qy 709 GTAGAAGAGTTACAAAATTTGAATGCTAGTTTCTTAATCTACTACAACTTTAGCTGGTAT 768
Db 661 GTAGAAGAGTTACAAAATTTGAATGCTAGTTTCTTAATCTACTACAACTTTAGCTGGTAT 720
Qy 769 GACAGAAATCAAAACAGCTATAGAGATAAGTAAAGAAATATTAACAATATGATGCGAGAAA 828
Db 721 GACAGAAATCAAAACAGCTATAGAGATAAGTAAAGAAATATTAACAATATGATGCGAGAAA 780
Qy 829 TCAGATCATTCAGCTGATGTTAAAGAGAAATGTTAAAGAAATGTTGATTTAGTAGTGCAAT 888
Db 781 CAAGATCATTCAGCTGATGTTAAAGAGAAATGTTAAAGAAATGTTGATTTAGTAGTGCAAT 840
Qy 889 GCACTAGTAGATGATTTGCGGCTCCTTTAGCAGCAGAAAAGATGCTCCACTATTA 948
Db 841 GCACTAGTAGATGATTTGCGGCTCCTTTAGCAGCAGAAAAGATGCTCCACTATTA 900
Qy 949 TTAACCTTCAAAAGATATAATTTAGATTCGTGAGTAAATCTGAAATAAGAGAGTTTAGAC 1008
Db 901 TTAACCTTCAAAAGATATAATTTAGATTCGTGAGTAAATCTGAAATAAGAGAGTTTAGAC 960
Qy 1009 TTAACCTTCAAAAGATATAATTTAGATTCGTGAGTAAATCTGAAATAAGAGAGTTTAGAC 1068
Db 961 TTAACCTTCAAAAGATATAATTTAGATTCGTGAGTAAATCTGAAATAAGAGAGTTTAGAC 1020
Qy 1069 GTATCTAAAGAGTTTGAACAGAAATTAAGATCAATGGGATTTAAAGTTGAAAGATTTCTCA 1128
Db 1021 GTATCTAAAGAGTTTGAACAGAAATTAAGATCAATGGGATTTAAAGTTGAAAGATTTCTCA 1080
Qy 1129 GGTGATGATAGATATGAACCTTCTTTTAAAGATAGCAGGTGAAATAGCTTTAGATAATGAT 1188
Db 1081 GGTGATGATAGATATGAACCTTCTTTTAAAGATAGCAGATGAAATAGCTTTAGATAATGAT 1140
Qy 1189 AAGGCTTATGATGTTGGTGGACAGATTTAGCAGATGCCATGAT 1233
Db 1141 AAGGCTTATGATGTTGGTGGACAGATTTAGCAGATGCCATGAT 1185

RESULT 5

AY004256
LOCUS 2434 bp DNA linear BCT 09-APR-2001
DEFINITION Clostridium difficile *slpA* (slpA) gene, complete cds.
ACCESSION AY004256
VERSION AY004256.1 GI:11496149
KEYWORDS Clostridium difficile
SOURCE Clostridium difficile
ORGANISM Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE 1 (bases 1 to 2434)
AUTHORS Karjalainen, T., Waligora-Dupriet, A.J., Cerquetti, M., Spigaglia, P.,
Maggioli, A., Mauri, P. and Mastrantonio, P.
TITLE Molecular and genomic analysis of genes encoding surface-anchored
proteins from Clostridium difficile
JOURNAL Infect.-Immun.-69-(5)-3442-3446-(2001)
MEDLINE 11292772
PUBMED 2 (bases 1 to 2434)
REFERENCE Karjalainen, T.K.

← noted Dec. 2006
AY004256
Clostridium difficile
slpA (slpA) gene, complete cds.
GI:11496149
Clostridium difficile
Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
1 (bases 1 to 2434)
Karjalainen, T., Waligora-Dupriet, A.J., Cerquetti, M., Spigaglia, P.,
Maggioli, A., Mauri, P. and Mastrantonio, P.
Molecular and genomic analysis of genes encoding surface-anchored
proteins from Clostridium difficile
Infect.-Immun.-69-(5)-3442-3446-(2001)
11292772
2 (bases 1 to 2434)
Karjalainen, T.K.

TITLE Direct Submission
JOURNAL Submitted (05-JUL-2000) Microbiology, Univ Paris-Sud, Faculty of Pharmacy, 5, rue JB Clement, Chatenay-Malabry 92296, France
FEATURES Location/Qualifiers

source

1. .2434
/organism="Clostridium difficile"
/mol_type="genomic DNA"
/strain="79685"
/db_xref="taxon:1496"
56. .2209
/gene="slpA"
56. .2209
/gene="slpA"
/note="S layer protein"
/codon_start=1
/transl_table=11
/product="SlpA"
/protein_id="AAF89093.1"
/db_xref="GI:11496150"
/translation="MSGLVLSAAPVPAADVKAEEYITVQDYKDTLKKIQAGIKDGS
ITNLVVTYDKREAVNANYKSDATTTADAKIEIAITLNLVDSKLNLGGDLVSNFIK
YDAEKFHKDEMDALTKLENKEIVKPASSETAGLVMADGATDSKADSKSYAKDVI
KFDVSDTTGYKLTATPIADAQLATKATKYANNNTKVFASATELAAIDGSAVEYAK
GKEYNATGSLVFSATGKTSNINVDPLNKGDTVKVINKESITDIDISSTGSAEDL
AKKYVFDEKLDLDIYKELTSEEGYNLVOLVSGRYOVALYPEGKRLDTKGATDIENTP
VKLVLRADKIKDKMDYIDDLRTYNNYSNVTVVAGEDRIETALIELSNKYNSDDKHAL
TSATDSVVLVGSQALVDGLVSLPLASEKHPALLLTSKKLDSNVKSEIKRVMDLKST
SGINTSKKVLVLAGVNSISKVENELDKMLKVLRLSGDDRYETSLIAIDEVGLNDK
AFVVGGTGLADAMSIAPVASQLKSNGLDLDVVDGATPIVVVDGKAKTINNETEDFLN
NAQVDIIGGNSVSKDVEKSIIVVATGKSPNRTSGDDRQATNAEVKMETDYFFBKGSVIN
YFVAKDGTSTKEDQLDALAAAPVAANFSGTYDGKNANGTVSPAPILVATDSLSDQNV
GYSKSVSDDGGKRLVQVGKLIASSVSKMKLLDM"

ORIGIN

Query Match 29.9%; Score 546.6; DB 1; Length 2434;
Best Local Similarity 59.0%; Pred. No. 2.3e-65;
Matches 1061; Conservative 0; Mismatches 684; Indels 54; Gaps 5;
QY 35 CTGTTACTGTAGTGGTTCTGCTGCTCCAGTCTTTTCGACGAGCGTTCAGATGTAATATCAC 94
DB 459 CAGCTGCTGTAGTAATGGCTGTGGTCTACTGATTCAAAAAAGCAGATAAAAAGTTTAT 518
QY 95 TACAAGATGTTACAAATGATAAGTATACAGTATCAAACTACTAAAGCTAGTGACTTAGTAA 154
DB 519 ATGCTAAGATGTAATTAAGTTTGATGTAGTATCAGATCTATAGATATAGATTTACAG 578
QY 155 AGGATATTTTACGACGACAAACCTTAACACAGGTGCAGTTATTTTGAACAAAGATACAA 214
DB 579 CAACACCAATTGCTGATGCACAATTAGCAACTTTAAAGCTACTTTATAAATATGCAATA 638
QY 215 AAGTTACTTCTATGATGCAAAATGAGAAAGATTCTTCAACTCOACTGGAGATAAAAAAG 274
DB 639 ATACAAAGGTGAGAAATTTGCTAGTGCTACAGAGCTTGCTGCTACAGATGGTTTCAGCTGTAG 698
QY 275 TTTTATTACAGAACAACTTTAACTACAGCTAATCGAATGAAGATTATGTAAGACAACTT 334
DB 699 AAGTTGCAAAAGGAAAGAAATATAAATCGGACTGGCTCACTGTTATTTGATGTGCTACTG 758
QY 335 TAAAAAATTTAGATGCAGGAGAAATATGCTATTATAGATTTAACTTATAAATATGCTAAAA 394
DB 759 GAAAGACTTCTAATATAAATGTTGATCCTTCTTAAACCAAGGGTGATACAGTTGTAAGAAG 818
QY 395 CTGTTGMAATTAAGTAGTAGAGCTAGTGAAAAAACAAGTAGTGTGTTATCTAGTGATGCGA 454
DB 819 TTAATAAATGCAAAAGAAATCAACAATAGATATAGATTCAAGTACTAGTACAAGTGTGGAAG 878
QY 455 AAAAATAGTCAAAAGAGATAGAGTCAAAATAATGTTTGAAGACAAAGACTTAGAAAAATG 514
DB 879 ATTTAGCTAAGAAATATGTTATTTGATGAAGATAAATTAGATGATATATATAAAGATTAA 938
QY 515 CACTAAAAACTATAAATGCCCTCAGATTTTCAGTAAAACTGATAGTTACTATCAAGTAGTTTC 574
DB 939 CTAGTGAAGAGGGATATGGAAATTTTAGTACAACACTAGTAGTGGGTAGATATCAAGTAGCTC 998

QY 575 TTTATCCAAAGGAAGAGATTACAAAGGTTTCTCAACTTATAGAGCTACAAATTTAATAG 634
DB 999 TTTATCCAAAGGAAGAGATT-----AGATACTAAAG 1031
QY 635 AAGGAACTGCATATCGTAAATACACAGATAATTAACCTCTAAAACTCTACTAGTAAGAGTA 694
DB 1032 GAGCAACAGATATAGAAAATACTCTCTGTTAAATTAGTTCTTAAAGCAGATAGATAAAG 1091
QY 695 ATTTAAAGATGTCAGTACAGAGATTACAAAATTTGAATGCTAGTTATTCTTAATATACAA 754
DB 1092 ATATCAAAAGATTATATAGATGATTATTAAGAACATATAACAACAGTATTTCAAATGTTGTA 1151
QY 755 CTTTAGCTGGTGATGACACAGATACAAAACAGCTATAGAGATAGTAAAGATATTAACAATA 814
DB 1152 CTGTAGCTGGAGAGATAGATAGAAACAGCTATAGAATTAGTAAATTAATATTAATTAAT 1211
QY 815 ATGATGGCGAGAAAATCAGATCAATTCAGCTGATGTTTAAAGAGAAATGTTAAAAATGTTGTAT 874
DB 1212 CTGAT-----GATAAGCATGCTATAAAGCTGATTCAGCAACTGATTCAGTAGTTT 1259
QY 875 TAGTAGGTGCAAAATGCATAGTAGATGATAGTTTGGCGCTCCTTTAGCAGCAGAAAAAG 934
DB 1260 TAGTTGGTCTCAAGCTATAGTTGATGCTCTTGTTGCATCACCTTTAGCATCAGAAAAAC 1319
QY 935 ATGCTCCACTATATTAATCTTCAAAAGATAAATTAGATTTCGTCACTAAATCTGAAATAA 994
DB 1320 ATGCTCCCATATATTAATCTTCAAAAGATAAACTAGATTCAAAATGTTAAATCTGAGATAA 1379
QY 995 ACAGAGATTTTACACTTAAAAAATCTCAACAGAGTA---ACAGGAAAAACAGTTTATATAG 1051
DB 1380 AAAGAGTTATGATTTAAAAATCTACAAGTGGTATAAATACTTCTTAAAAAGTTTATTTAG 1439
QY 1052 CTGGTGGAGTTAATAGTGTATCTAAAGAGTTGTTAAACAGATTAAGATCAATCGGATTA 1111
DB 1440 CTGGTGGAGTTAACTCTATATCTAAAGAGTTGAAATCAATTAAGAGATATGGGACTTA 1499
QY 1112 AAGTTGAAGAGTCTCAGGTGATGATAGTATGAACTCTTTAAAAATAGCAGCTGAAA 1171
DB 1500 AAGTTACAAGATATCTCGTGATGATGAGATATGAACTCTTTGCTATAGCTGATGATGAG 1559
QY 1172 TAGGCTTAGATAATGATAAGGCTTATAGTATGGTGGAAACAGGATTAGCAGATGCCATGA 1231
DB 1560 TAGGCTTGTGATGATGATAAGCAATTTGATGTTGGTGGAACTGGATTAGCAGATGCTATGA 1619
QY 1232 GTATAGCTTCAGTTGCTTCTACTAAATTAGATGATGTTAAATGTTGTTGTTGATAGATAAAGATG 1291
DB 1620 GTATAGCTCCAGTTGCTTCTCAATTAAGAGAAATCTAATGGAGATTTAGATGTAGTTGATG 1679
QY 1292 GACATGCTACTCCAATAGTTGTTGATGAGTGGAAAGCTGATAAATAATCTGATGACTTAG 1351
DB 1680 GAGATGCTACTCCAATAGTTGTTGATGAGTGGAAAGAGCTAAAACTATAAATAATGAACAG 1739
QY 1352 ATAGTTTCTTAGGAAAGCGCTGATGTAGATATAATAGTGGGATTTGCAAGTGTATCTGAAA 1411
DB 1740 AAGATTTCTTAAACAATGCAACAAGTTGATATAATAGTGGGAGAAACAAGTGTATCTAAG 1799
QY 1412 AGATGGAAGAAGCTATATCAGATGCTACTGTTAAAGGGGCTTACAAGATTTAAAGGCGAG 1471
DB 1800 ATGTAGAAAAAATCTATAGTTGTTGCTACTGGAAGAGAGCCTAACAGAACTAGTGGAGATG 1859
QY 1472 ATAGACACAGACTCAACTCTGAAGTTATAAACAATATTTGCTTAATGATACTGAATAAG 1531
DB 1860 ATAGACAAAGCTACTAATGCAAGATTAATGAAGAAACTGATTAATCTTCAAAAAAGGAGTG 1919
QY 1532 CTAAAGCTGCAGTTTTAGATAAAGATTTCAGGTGCTTCAAGTAGTAGTATGATGACGAGTATTTTA 1591
DB 1920 TAATAAACTACTCTCGTAGCAAAAGAT---GGTTCTACTTAAAGAGATCAATTTAGTAGATG 1976
QY 1592 ATTTCTATGTAGCTTAAAGATGATCTACAAAAGAGATCAATTAGTTGATGATTAAGCAG 1651
DB 1977 CTTTAGCAGCAGCTCCAGTTGAGCTAACTTCGGTTCTACTTATGATGTTGTAATAAAGCGCTA 2036

QY 1652 TAGGAGCTGTTGGATATAAATCTTCCAGTGTGTTAGTACTGATCTTTATCTT 1711
 Db 2037 ATGGAACTGTTCT-----CCAGCTCCATAGTATTAGTACTGATAGTATCTG 2087
 QY 1712 CTGATCAATCGGTTGCTATAAGCAAAAGTTGTAGGAGAAAATATCTTAAAGATTAAAC 1771
 Db 2088 CAGACCAAAATGTAGGTGAAGTAAATCAGTTAGTGATGTGTGGAAGAACTTAGTTTC 2147
 QY 1772 AAGTTGGTCAAGAAATAGCTAAATTCAGTTATAACAAATGAAAGATTTTATTAGATATG 1830
 Db 2148 AAGTTGGTAAAGGTATCGCTAGTTTCAGTAAATAGCAAAATGAAAGATTTTATTAGATATG 2206

RESULT 6
 AX512904 2271 bp DNA linear PAT 03-OCT-2002
 LOCUS
 DEFINITION Sequence 4 from Patent WO02062379.
 AX512904
 ACCESSION
 VERSION AX512904.1 GI:23504063
 KEYWORDS
 SOURCE Clostridium difficile
 ORGANISM Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

REFERENCE
 1 Doyle, R., Kelleher, D., Windle, H. J., Walsh, J. B. and Deirdre, N. B.
 Clostridium difficile vaccine
 Patent: WO-02062379-A 4 15-AUG-2002;
 THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH (IE)

FEATURES
 source Location/Qualifiers
 1..2271
 /organism="Clostridium difficile"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1496"

ORIGIN
 Query Match 29.6%; Score 542; DB 6; Length 2271;
 Best Local Similarity 59.4%; Pred. No. 9.8e-65;
 Matches 1117; Conservative 0; Mismatches 670; Indels 93; Gaps 8;

QY 32 CAGCTGTTACTGTAGTAGTCTCTGCTCCAGTCTTTTGCAGCAGCTTCAGATGTAATAT 91
 Db 401 CAGCAGATGCTAATATTTGCTGGACATCTTCAGCAGATGGTGTGTTACAAATCTGGAG 460
 QY 92 CACTACAAGATGGTACAAATGATAGTATACAGTATCAAAATCTAAAGCTAGTACTTAG 151
 Db 461 CTGCTAGTGGTTCTACTGAGACAAATTCAGCAGGACAAACTTGCATGTCTAGCTATTT 520
 QY 152 TAAAGGATATTTTAGCAGCAGCAAACTTAACACAGGTGCGATTTATTTTGAACAAAGATA 211
 Db 521 TTGACACAGCATATACAGATTCATCTGAAACTGCGGTAAAGATTACTATATAAAGCAGATA 580
 QY 212 CAAAGTTACT---TTCTATGATCAATGAGAAGATCTTCAACTCCAACTCGAGATA 268
 Db 581 TGAATGATACATAAATTTGGTAAAGCAGGTGAGACACTTATTAATCTGGGCTTACATTG 640
 QY 269 AAAAAGTTTATTCAGAACAAACTTTAATACAGCTAATGGAATGGAATGGAATGATGTAAGA 328
 Db 641 AAGATGGCTCTACAGAAAATTTGTAATTTAGGGGAGTGTATATTATAGATATACTA 700
 QY 329 CAACCTTTAAA---AATTAGATGAGGAGAAATATGCTATTATATAGATTTAACTTATATA 385
 Db 701 AAGCTCTTAAACTTACTGTTCTCTGGAAGTAAAGCAACTGTGTTAAGTTTGTCTGAAAAA 760
 QY 386 ATGCTAAAACGTGTTG-----AAATTAAGTAGTACAGCTAGTGA 427
 Db 761 CACCAAGTCCAGGTGTTCAACAGTAAATACAAAGCTTAGAATTAATTAATCTTAAGAG 820
 QY 428 AAACAGTAGTTGTATCTAGTATGCGAAAAATATAGTGCAAAAAGATATATAGCTGAAAAATATG 487
 Db 821 AAACATAGATATTGACCTAGTTCTAGTAAACAGCAGCAAGATTTAGCTAAAAAATATG 880

QY 488 TGTTTGAAGACAAAGACCTTTAGAAAATGCACTAAAACTATAAATGCCTCAGATTTCACTA 547
 Db 881 TATTTAAATAAACTGATTTAAATCTCTTTATAAAAGTATTAATAATGAGATGAAGCAGATA 940
 QY 548 AAACGTAGATTACTATCAAGTAGTCTTTTATCCAAAAGAAAGAGATTACAAGGTTTCT 607
 Db 941 CTAATGGATTAATAGAGAAGTTAGTGGAAAATATCAAGTAGTTCTTTATCCAGAAGGAA 1000
 QY 608 CAACTTTATAGACTACAAATTTAT---AATGAAGGAACCTGCATATATGTTAAATACACAGTAA 664
 Db 1001 AAGAGTTACACTAAGAGTCTGCAAGGCTTCAATTCGTGATGAAATTCACCAAGTTA 1060
 QY 665 TATTAACCTCTAAAATCTACTAGTAAGAGTAATTTTAAAGACTGCGAGTAGAAGATTACAA 724
 Db 1061 AATTAACCTCTTAAGTCAGATAAGAGCAAGAGACTTAAAGATTTATGTTGGATGATTTAAGAA 1120
 QY 725 AATTGATGCTAGTTATTTCTTAATCTACTACAACTTTAGCTGCTGATGACAGATAACAACAG 784
 Db 1121 CATATAAATAGGATATTCARATGCTATAGAGTAGCAGAGAGAGATAGATAGAACTG 1180
 QY 785 CTATAGAGATAAGTAAAGAAATATTACAATAATGATGCGAGAAATCAGATCATTTACGCTG 844
 Db 1181 CAATAGCATTAAGTCAAAAATATTATAA CTCTGATGATGAAA-----TGCTA 1228
 QY 845 ATGTTAAAGAGAAATGTTAAAAATGTTGTATTTAGTAGGTGCAATGCACTAGTAGATGGAT 904
 Db 1229 TATTTAGAGATTTCAGTTGATATGTTAGTTGTTGGAGAAATGCAATAGTTGATGGAC 1288
 QY 905 TAGTTGGGCTCTTTAGCAGCAGAAAAGAGTCTCCACTATTATTAACTTCAAAAGATA 964
 Db 1289 TTGTAGCTTCTCCTTTAGCTTCTGAAAAGAAAGCTCTCTTTATTATTAACTTCAAAAGATA 1348
 QY 965 AATTAGATTGCTCAGTAAATCTGAAATAAGAGAGTTTTTAGACTTAAAAAATTTCAACAG 1024
 Db 1349 AATTAGATTCAACGGTAAAGCTGAAATAAAGAGAGTTATGAATATAAAGAGTACAACAG 1408
 QY 1025 AAGTA---ACAGAAAAACAGTTTATATAGCTCGTGGAGTTAATAGTGTATCTAAAGAAG 1081
 Db 1409 GTATAAATCTCTCAAGAAAGTTTATTTAGCTGCTGGAGTTAATTTCTATATCTAAAGAG 1468
 QY 1082 TTGTAACAGAAATTAGAAATCAATGGGATTAAGGTTGAAAGATTTCTCAGGTGATGATAGAT 1141
 Db 1469 TAGAAAAATGAATTTAAAGATATGGGACTTAAAGTTTACAAGATTAGCAGGAGATGATAGAT 1528
 QY 1142 ATCAAACTTCTTTAAAAATAAGCAGGTGAAATAGCTTTAGATAATGATAAGGCTTATCTAG 1201
 Db 1529 ATGAACTTCTCTAAAAATAGCTGTAGAGTAGCTTTGATATATGATAAGCATTTCTAG 1588
 QY 1202 TTGTTGGAACAGGATTAGCAGATGCCATGATAGTATAGCTCCAGTTGCTTCTCAATTAAGAA 1261
 Db 1589 TTGGAGGAACAGGNTTAGCAGATGCCATGATAGTATAGCTCCAGTTGCTTCTCAATTAAGAA 1648
 QY 1262 ATGTTAATGTTGTTAGATAGAACAAATGGACATGCTACTCCNATAGTTGTTGTTAGATG 1321
 Db 1649 ATGCTAATGTAATAATGGATTTAGCTGATGCTGATGCTACCAATATAGTATGTTAGATG 1708
 QY 1322 GAAAGCTGATAAATCTGATGACTTTAGTATGTTTCTTAGGAAGCGCTGATCTAGATA 1381
 Db 1709 GAAAGCTAATACTATAAATGATGATAAAGATTTCTTAGATGATTCAAGTTGATA 1768
 QY 1382 TAAATAGTGGATTTGCAAGTGTATCTGAAAAGATGGAAGAGCTTATATCAGATGCTACTG 1441
 Db 1769 TAAATAGTGGAGAAAACAGTGTATCTAAAGATGTTGAAAATGCAATAGATGATGCTACAG 1828
 QY 1442 GTAAAGCGGTTACAAGAGTTTAAAGCGCAGATAGACAAAGACATCACTCTGGAAGTTATA 1501
 Db 1829 GTAATCTCCAGATAGATATAGTGGAGATGATAGACAAGCAACTAATTCGAAAAGTTATA 1888
 QY 1502 AAACATATTATGCTAATGATACTGAAATAGCTAAAGCTGCAAGTTTATAGATAAAGATTCCAG 1561
 Db 1889 AAGAATCTTCTTATTATCAAGATAACTTAAATATGATATAAATAAGTAGTTTATTTCTTTG 1948
 QY 1562 GTGCTTCAAGTAGTAGTGACGAGTATTTAATTTCTATGTAGCTAAAGATGGATCTACAA 1621

1121 CATATATATGATATTCATATGCTATAGAGTAGCAGGAGAGATAGAAATAGAACTG 1180
 785 CTATAGAGATAGTAAGAAATATATCAATATAGTATGCGAGAGAAATCAGATCATTCAGCTG 844
 1181 CAATAGCATTAAGTCAAAATATATTAATCTGATGATGAAAA-----TGCTA 1228
 845 ATGTTAAAGAGATGTTAAAAATGTTGTTATTTAGTGGTCAAAATGCACCTAGTAGATGAT 904
 1229 TATTTAGAGATTCAGTTGATATGTTAGTATTTGTTGGAGAAATGCAATAGTTGATGAC 1288
 905 TAGTTGCGGCTCCTTTAGCAGCAGAAAAAGATGCTCCACTATATTAATTAATCAAAAGATA 964
 1289 TTGTAGCTTCTCTTTAGCTTCTGAAAGAGAAAGCTCTTTATTTATTAATCTCAAAAGATA 1348
 965 AATTAGATTCGTAGTAAATCTGAAATTAAGAGAGTTTGTAGACTTAAAAATCTCAACAG 1024
 1349 AATTAGATTCAGCGTAAAGCTGAAATTAAGAGAGTTTGTAGATTAATTAAGAGTACAAACAG 1408
 1025 AAGTA---ACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGTATCTAAAGAG 1081
 1409 GTATAAATCTTCAAAGAAAGTTTATTTAGCTGGTGGAGTTAATCTATATCTTAAGAG 1468
 1082 TTGTAAACAGATTAAGTCAATGGGATTAAGAGTTGAAAGATTCCTCAGGTGATGATGAT 1141
 1469 TAGAAATGAATTAAGATATGGGACTTAAAGTTAAGATTTAGCAGCAGATGATGAT 1528
 1142 ATGAAACTCTTTAAAAATAGCAGGTGAATAGGCTTAGATTAATGATAGGCTTATGATG 1201
 1529 ATGAAACTCTCTAAAAATAGCTGATGAAGTAGGCTTTGTATATGATTAAGAGCAATTTGTAG 1588
 1202 TTGGTGGACAGATTTAGCAGATGCCATGATAGTATCTCAGTTGCTTCTACTAAATTAG 1261
 1589 TTGAGGAAACAGATTTAGCAGATGCCATGATAGTATCTCAGTTGCTTCTACTAAATTAG 1648
 1262 ATGGTAATGGTGTGTAGATAGAACAAATGGACATGCTACTCCAAATAGTTGTTGTAGATG 1321
 1649 ATGCTAATGGTAAATGGAATTTAGCTGATGGTGTGCTACACCAATAGTAGTTGTAGATG 1708
 1322 GAAAGCTGATAAAATATCTGATGATGATAGTATGATGTTCTTGAAGCGGTGATGATGATA 1381
 1709 GAAAGCTGATAAACTATAATGATGATGATGATGATGATGATGATGATGATGATGATGATA 1768
 1382 TAATAGGTGATTTGCAAGTGTATCTGAAAGATGGAAGATGATGATGATGATGATGATGATG 1441
 1769 TAATAGGTGGAAGAAACAGTGTATCTAAAGATGTTGAAATGCAATGATGATGATGATGATG 1828
 1442 GTAAGGCGTTACAAGAGTTAAAGGCGACGATAGACAAAGACACTTAATCTGAAAGTTATAA 1501
 1829 GTAAATCTCAGATAGATATAGTGGAGATGATAGACAAAGCACTTAATGCAAAAGTTATAA 1888
 1502 AAACATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1561
 1889 AAGAACTCTTCTTATTAATCAAGATAAATCTTAATAATGATTAATAAAGTAGTTAATTTCTTTG 1948
 1562 GTGCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1621
 1949 TAGCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2008
 1622 A-----AGAAATCAATTAATGATG 1639
 2009 TTGAGCAAACTTTGGTGAATCTTAATTTCTGATGTTAGCCAGTAGATTAAGATGTTA 2068
 1640 ATGCATTTAGCAGTAGGAGCTGTTGCTCGATATAA-----CTTGCTCCAGTTGTTAT 1690
 2069 AAGTATTAATCTGTTCTGATATAATGATAAATAAATTAATTTAGTATCTCCAGCACCTATAGTAT 2128
 1691 TAGCTACTGATTTCTTCTGATCAATCGGTTGCTATTAAGCAAAAGTTGTAGGAGAAA 1750
 2129 TAGCTACTGATTTCTTCTGATCAATCGGTTGCTATTAAGCAAAAGTTGTAGGAGAAA 2188
 1751 AATATTTCAAGATTTAACAACAGTTGCTCAAGGAATAGCTTAATTTCAAGTTATTAACAAA 1810

Db 2189 ATAATGGAGAAAACCTTAGTTCAGGTGGTAAAGGTATAGTACTTCAAGTTATAAACAAT 2248
 Qy 1811 TCAAAAGATTTATAGATATG 1830
 Db 2249 TAAAGATTTATTAAGTATG 2268
 RESULT 8
 AX259458 2271 bp DNA linear PAT 26-OCT-2001
 LOCUS Sequence 6 from Patent WO0173040.
 DEFINITION
 ACCESSION AX259458
 VERSION AX259458.1 GI:16508626
 KEYWORDS
 SOURCE Clostridium difficile
 ORGANISM Clostridium difficile
 Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium.
 1
 REFERENCE
 AUTHORS Fairweather, N.F. and Calabi, B.D.
 TITLE Clostridium difficile polypeptides and uses thereof
 JOURNAL Patent: WO 0173040-A 6 04-OCT-2001;
 IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY & MEDICINE (GB)
 FEATURES
 source
 1..2271
 /organism="Clostridium difficile"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1496"
 ORIGIN
 Query Match 29.3%; Score 535.6; DB 6; Length 2271;
 Best Local Similarity 59.2%; Pred. No. 7.3e-64;
 Matches 1113; Conservative 0; Mismatches 674; Indels 93; Gaps 8;
 Qy 32 CAGCTGTTACTAGTAGGTTCTGCTGCTCCAGTTTTCGACGAGCTTCAGATGTAATAT 91
 Db 401 CAGCAGATGCTAATTTGCTGGAACATCTTCAGCAGATGGTGTGTTTACAAATCTCGGAG 460
 Qy 92 CACTACAAGATGGTACAAATGATAGTATACAGTATCAAACTACTAAAGCTAGTGACTTAG 151
 Db 461 CTGCTAGTGGTCTACTGAGACAAATTCAGCAGGACAACTTGCATGTCAGCTATTT 520
 Qy 152 TAAAGGATATTTAGCAGCACAACCTTAACAAAGGTGCGAGTATTTTGAACAAAGATA 211
 Db 521 TTGACACAGCATATACAGATTCATCTGAAACTCGCGTTAAAGATTACTATATAAAGCAGATA 580
 Qy 212 CAAAAGTTACT---TTCTATGATGCAATGAGAAAGATTTCTCAACTCCAACTGGAGATA 268
 Db 581 TGAATGATATAAATTTGGTAAAGCAGGTGAGACAACTTAATCACTGGGCTTACATTTG 640
 Qy 269 AAAAACTTTATTACAGAAACAACTTTTAACCTACAGCTAATGGAATGAAGATTATGTAAGA 328
 Db 641 AGATGGGTCTACAGAAAAAATTTGTTAAATTAGGGACAGTGATATATAGATATACTA 700
 Qy 329 CAACTTTAAAA---AATTTAGATGAGGAGAAATATGCTATTTATAGATTTTAACTTATATA 385
 Db 701 AAGCTCTTAACTTACTGTTGTTCTCTGGAAGTAAAGCAACTGTAAAGTTTGTCTGAAAAA 760
 Qy 386 ATGCTAAAACCTGTTG-----AAATTAAGTAGTAGCAGCTAGTGAAA 427
 Db 761 CACCAAGTGCAGTGTTCACCCAGTAATAACAAAGCTTGAATAATAAATGCTAAAGAG 820
 Qy 428 AACACGTAGTTGATCTAGTATGCGAAAAAATAGTCAAAAGATATAGCTGAAAAATATG 487
 Db 821 AAACAATAGATATGACGTAGTTCTTAGTAAACAGCAAGATTTAGCTTAAAAATATG 880
 Qy 488 TGTTTGAAGCAAAAGACTTTAGAAAAATGCACTAAAAAATAAATGCTCAGATTTTCAGTA 547
 Db 881 TATTTAATAAACTGATTTAAATACCTTTTATTAAGTATTTAAATGGAGATGAAGCAGATA 940
 Qy 548 AAACGTAGTATCTATCAAGTAGTCTTTTATCCAAAGGAAAGAGATTACAGGTTTCT 607
 Db 941 CTAATGATTAATAGAGAGTAGTTAGTGGAAAAATATCAAGTAGTTCTTTTATCCAGAGAA 1000

```
QY 608 CAACTTATAGAGCTCAAAATTAT---AATGAAGGAACTGCATATGCTATATACACAGTAA 664
Db 1001 AAGAGTTCAACTAAGAGTGCCTCAAAAGCTTCAATTGCTGATGAAATTCACCAAGTTA 1060
QY 665 TATTAACCTCAAAATCTACTAGTAGAGTAATTTAAAGAACTGCAGTAGAAGAGTTACAAA 724
Db 1061 AATTAACCTTAAAGTCAGATAGAGAAAGACTTAAAGAGTTATGTGGATGATTTAAGNA 1120
QY 725 AATTGAATGCTAGTTATTTCTAATACACAACTTTAGCTGGTGATGACAGAATACAAACAG 784
Db 1121 CATATAAATAATGGATATTCAAATGCTATAGAAAGTAGCAGGAGAGATAGAAATAGAACTG 1180
QY 785 CTATAGAGTAAGTAAGAAATATACAAATGATGGCGAGAAATCAGATCATTCAGCTG 844
Db 1181 CAATAGCATTTAAGTCAAAATATTTAACTCTGTGATGATGAAAA-----TGCTA 1228
QY 845 ATGTTAAAGAGAATGTTAAAAATGTTGATTAGTAGTGCAAAATGCACCTAGTAGATGGAT 904
Db 1229 TATTTAGAGATTCAGTTGTATATGTAGTATGTTGGAGGAATGCAATAGTTGATGNC 1288
QY 905 TAGTTGCGGCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTTATTAACCTTCAAAAGATA 964
Db 1289 TTGTAGCTTCTCTTTAGCTTCTGAAAAGAAAGCTCCTTTATTTAACTTCAAAAGATA 1348
QY 965 AATTAGATTCGTCAGTAAATCTGAAATAGAGAGATTTTAGACTTAAAACTTCAACAG 1024
Db 1349 AATTAGATTCAGCGCTAAAGCTGAAATTAAGAGAGTTATGAATATAAGAGTACACAG 1408
QY 1025 AAGTA---ACAGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGTATCTAAAGAG 1081
Db 1409 GTATAAATACTTCAAGAAAGTTTATTTAGCTGGTGGAGTTAATCTATATCTAAAGAG 1468
QY 1082 TTGTAACAGAAATTAAGATCGGAATTAAGATTTGAAAGATTCCTCAGTGATGATAGAT 1141
Db 1469 TAGAAAAATGAAATTAAGATATGGGACTTAAAGTTTACAGATTTAGCAGGAGATGATAG 1528
QY 1142 ATGAAATCTTTTAAATATAGCAGGTGAATAGGCTTAGATTAATGATAGGCTTATAGTAG 1201
Db 1529 ATGAAATCTTCTTAAATATAGCTGATGATGAAATGAGTGGTCTTTGATTAATGATTAAG 1588
QY 1202 TTGTTGGAAACAGGATTTAGCAGATGCCATGAGTATAGCTTCAGTTGCTTCTACTAAATAG 1261
Db 1589 TTGGAGGAACAGGATTTAGCAGATGCCATGAGTATAGCTTCCAGTTGCAATCTCAATTAAG 1648
QY 1262 ATGTTAATGGTGTGTAGATAGAAACAAATGGACATGCTACTCCAATAGTTGTTGATAGT 1321
Db 1649 ATGCTAAATGGTAAATGGATTTAGCTGATGGTGTGCTACACCAATAGTTGTTGATAGT 1708
QY 1322 GAAAGCTGATAAAATATCTGATGACTTAGATAGTTTCTTAGGAAGCGCTGATGATAGATA 1381
Db 1709 GAAAGCTTAAACTATAATGATGATGTTTAAAGATTTCTTAGATGATTCACAAAGTTGATA 1768
QY 1382 TAATAGGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATCAGATGCTACTG 1441
Db 1769 TAATAGGTGGAGAAACAGTGTATCTAAGATGTTGAAATGCAATAGATGATGCTACAG 1828
QY 1442 GTPAAGGGGTTACAGAGTTAAAGGGCAGATAGACAAGACACTAACTCTGAAAGTTATAA 1501
Db 1829 GTAAATCTCCAGATAGATAGTGAGATGATAGACAAGCAACTAATGCAAAAGTTATAA 1888
QY 1502 AAACATATATCTAATGATACCTGAATAGCTTAAAGCTGCAGTTTATAGATAAGATTCAG 1561
Db 1889 AAGAATCTCTTTATATCAAGATACTTAAATGAATGATAAAAGTAGTTAATTTCTTTG 1948
QY 1562 GTGCTTCAAGTAGTGTAGCAGGATTTTAAATTTCTATGTAGCTAAAGATGATCTACAA 1621
Db 1949 TAGCTAAAGATGGTTCTACTAAGAGAGATCAATTAGTTGATGCTTTTAGCAGCAGCTCCAG 2008
QY 1622 A-----AGAGATCAATTAGTTG 1639
Db 2009 TTGCAGCAAACTTTGGTGTAACTCTTAATTTCTGATGGTAAGCCAGTAGATAGATAAGATG 2068
```

```
QY 1640 ATGCATTAGACTAGGAGCTGTTGCTGATATAAA-----CTTGCTCCAGTTGTAT 1690
Db 2069 AAGTATTAACTGGTCTCGATATGATAAAATAAATATAGTATCTCCAGCACTATAGTAT 2128
QY 1691 TAGCTACTGATTTCTTTATCTCTGATCAATCGGTTGCTATAGCAAAAGTTGTAGAGAAA 1750
Db 2129 TAGCTACTGATTTCTTTATCTCTCAGATCAAAAGTGATCTATAGTAAGTTCTTGATAAG 2188
QY 1751 AATATCTTAAGATTTTAAACAAAGTTGGTCAAGGAATAGCTTAATTCAGTTTATAAACAANA 1810
Db 2189 ATAATGGAGAAAACCTTAGTTCAAGTTGGTAAAGGTATAGCTACTTTCAGTTTATAAACAAT 2248
QY 1811 TGAAAGATTTTATAGATATG 1830
Db 2249 TAAAGATTTTATTAAGTATG 2268

RESULT 9
AX512906 2217 bp DNA linear PAT 03-OCT-2002
LOCUS AX512906 Sequence 6 from Patent WO02062379.
DEFINITION AX512906
ACCESSION AX512906
VERSION AX512906.1 GI:23504065
KEYWORDS Clostridium difficile
SOURCE Clostridium difficile
ORGANISM Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
1
REFERENCE
AUTHORS Doyle R., Kelleher D., Windle H.J., Walsh J.B. and Deirdre N.E.
TITLE Clostridium difficile vaccine
JOURNAL Patent: WO 02062379-A 6 15-AUG-2002;
THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH (IE)
FEATURES
source Location/Qualifiers
1..2217
/organism="Clostridium difficile"
/mol_type="unassigned DNA"
/db_xref="taxon:1496"
ORIGIN
Query Match 27.5%; Score 503.4; DB 6; Length 2217;
Best Local Similarity 59.6%; Pred No. 1.8e-59;
Matches 950; Conservative 0; Mismatches 616; Indels 27; Gaps 5;
QY 250 TCAACTCCAAGTGGAGATAAAAAAGTTTATTTCAGAACAAACTTTTAACTACAGCTAATGGA 309
Db 640 TCTGTTCTCTACTACAGGCTTAACATTAACCTGCTGATACAACTGCAACAACAGATGTAAT 699
QY 310 AATGAAGATTTATGTAAGACAACTTTAAAAAATTTAGATGCGAGGAGATATGCTATTATA 369
Db 700 ATTTCTGATGTTATGAGTGCATTTAAATTTTAAATGCTACTGATACGATAGTGGATTCOCA 759
QY 370 GATTTAACTTATAATAAGCTGTAATAACTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAANA 429
Db 760 GCTGTTTCATCAGCTTCTACTCTTTAGACCAAGTATAAAGTAATTAATGCAAAAGAGNA 819
QY 430 ACAGTAGTTGATCTTAGTGATGCGAAAAATAGTGCAAAAGATATAGCTGGAATAATATGTG 489
Db 820 TCTATAGATGTTGATTCAAGTTTCAATAGAACAGCTGAAAGATTTAGCTGAAAAATATGTA 879
QY 490 TTTGAGACAAGACTTAGAANAATGCATAAAACTATAAATGCTCAGATTC----- 543
Db 880 TTTAAACAGAGAATGTTGAATAAAACCTTATGAGGCACCTGATGTTTATATAAAGAAGGT 939
QY 544 AGTAAACTGATAGTTACTATCAAGTAGTTCTTTTATCCAAAAGGAAGAGATTACAAGT 603
Db 940 ATACCAAGTATCTTATCACTCAAGATGGTGGAAATATCAAGTTGTTTATTTGCTCNA 999
QY 604 TTCTCAACTTATAGAGCTACAAATTAATGAAGGAATCTGCATATGGTAAATACACAGTA 663
Db 1000 GGAAGAGATTAACCTACTTAAGAGGCAACTGGAACTTTTAGCAGATGAAAAATTTCTCTCT 1059
```


664 ATATTAACTCTAAATCTACTAGTAGAGTAATTTAAAGACGTCAGTAGAAGAGTTACAA 723
1060 AAGGTAAACAATAAAGCAGATAAAGTAAAGACTTAAAGATATATGTTGAAGATTAAAA 1119
724 AAATTGAATGCTAGTTATTCTTAATACTACAACCTTTAGCTGGTGATGACAGAAATACAAACA 783
1120 AATGCTAACCAATGGATATTCAAAATCTGTGTTGTAGCAGGTGAAGATAGAAATAGAAACA 1179
784 GCTATAGAGATAGTAAAGAAATATTACAAATATATGATGGCGAGAAATCAGATCAATTCAGCT 843
1180 GCAATAGAGTTAAGTAGCAAAATCTATAAATCTCTGATGATGACAA-----TGCA 1227
844 GATGTTAAAGAGAACTGTTAAAGATGTTGATTAGTAGGTGCAATGCACTAGTAGATGGA 903
1228 ATAACCTAAAGATCCAGTTAAACATGTTGTTTATGTTGTTCTCAAGCTGATGATGGG 1287
904 TTAGTTGCGGCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAATCTCAAAAGAT 963
1288 CTGTAGCTTCACCTTTAGCATCTGAAAAAAGAGCTCCTTTACTATTAACTTCAGCAGGA 1347
964 AAATTAGATTGTCAGTAAATCTGAAATAAAGAGAGTTTATAGACTTAAAACTTCAACA 1023
1348 AAATTAGATTCAAGCTGTTAAAGCTGATTTGAAAGAGATTAATGGAATTTAAATCTACACA 1407
1024 GAAGTA---ACAGGAAAAAGCTTTATATAGCTGGTGGAGTTAATAGTGTATCTAAAGAA 1080
1408 GGTGTAATTAATCTCTAAAGAGTTTACTAGCTGGTGGAGTAACTCTATATCTAAGAT 1467
1081 GTTGTAAACAGAAATTAGAAATCAATGGGATTTAAAGTTTGAAGATTTCTCAGGTGATGATA 1140
1468 GTAGAAAAATGAATTAAGAGATATGGGACTTAAAGTTTACAAGATTATCAGGAGATGATA 1527
1141 TATGAAATCTTTTAAATATAGCAGGTGAATAGGCTTAGATTAATGATAGGCTTATGTA 1200
1528 TATGAAATCTTTTATAGCTATAGCTGATGAATAGGCTTGTGATTAATGATGAAGCTTTGTA 1587
1201 GTTGTGGACAGATTTAGCAGATGATCAGATATAGCTTCACTTCTCTCTAAATTA 1260
1588 GTTGGAGGACAGATTTAGCGGATGCTATGAGTATAGCTCCAGTTGCTTCTCAATTAAGA 1647
1261 GATGTAATGCTGTTGTAGATAGAAACAAATGGACATGCTACTCCAAATAGTTGTTAGAT 1320
1648 AACTCAATGGAGAACTTGA---CTTAAAGGTGATGCAACTCCAATAGTAGTTGTGAT 1704
1321 GGAAGAGCTGATATAATCTGATGATCTAGATGATTTCTTAGGAGCGCTGATGATAGAT 1380
1705 GGAAGAGCTAAAGATATAAATCTGAAAGTAAAGATTTCTTAGATGATTCACAAAGTTGAT 1764
1381 ATATAGGTGATTTGCAAGTGTATCTGAAAGATGCAAGATCTATCAGATGCTACT 1440
1765 ATATAGGTGATTTGAAATAGTGTCTTAAAGAGATTAATGGAAGCAATAGATGATGCTACT 1824
1441 GGTAAAGCGGTTACAAGAGTTAAAGCGGACGATAGACAGACACTACTCTGAAGTTATA 1500
1825 GGAATAATCACTGAGAGATATAGTGGAGAGATAGACAGCAACAAATGCTAAAGTTATA 1884
1501 AAAACATATTATGCTAATGATATCTGAAATAGCTAAAGCTGCGATTTAGATATAAGATTCA 1560
1885 AAGAAGATGATTTCTTTAAAGATGGAAGTTACAAACTTCTTTGTAGCTAAAGATGGT 1944
1561 GGTGCTTCAAGTAGTATGACAGGAGTATTTAAATTTCTATGATGCTAAAGATGATCTACA 1620
1945 TCAATAAAGAAAGATCAATTAATAGTAGCTTTAGCAGGTGCTGCAATGCTGGTAACTTT 2004
1621 AAAGAAGATCAATTAATAGTGTATGACAGTAGGAGCTGTTGCTGGATATAAATCTTGCT 1680
2005 GGTGTAACAGTAGATATAAGAGAAACCTACAGTTGCTGATATAAAGCTTCTCCAGCA 2064
1681 CCAGTTGATTAGCTACTGATCTTTATCTTCTGATCAATCGGTTGCTATAGCAAAAGTT 1740
2065 CCAATGTTTTAGCAACAGATCTTCTTATCTGATCAAAATGATGCTATATAAGTAAAGCT 2124
1741 GTA---GGAGAAAAATATTTCTAAGAGATTTAACAAGTTGCTGCAAGGATAGCTAATTC 1797

2125 GTAAATGATGCGCTAAATCTAAGAACTAGTCAAGTTGGTAAAGGTATAGCTACTTCA 2184
1798 GTTATAACAAAAATGAAAGATTTATTAGATATG 1830
2185 GTTGTAAATAAATAAAGATTTATTAGATATG 2217
CDI300677 2145 bp DNA linear BCT 04-SBP-2001
Clostridium difficile alp gene for S-layer protein, strain R7404.
AJ300677
S-layer protein; alp gene.
Clostridium difficile
Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
1
Calabi, E., Ward, S., Wren, B., Paxton, T., Panico, M., Morris, H.,
Dell, A., Dougan, G. and Fairweather, N.
Molecular characterization of the surface layer proteins from
Clostridium difficile
Mol. Microbiol. 40 (5), 1187-1199 (2001)
21295357
11401722
2 (bases 1 to 2145)
Fairweather, N.P.
Direct Submission
Submitted (19-DEC-2000) Fairweather N.F., Biochemistry, Imperial
College, Exhibition Road, London, SW7 2AY, UNITED KINGDOM
Location/Qualifiers
1. 2145
/organism="Clostridium difficile"
/mol_type="genomic DNA"
/strain="R7404 (ribotype #17)"
/db_xref="taxon:1496"
/country="United Kingdom"
1. 2145
/gene="alp"
1. 2145
/function="surface layer protein"
/evidence=experimental
/transl_table=11
/product="S-layer protein"
/protein_id="CAC35721.1"
/db_xref="GI:13539184"
/db_xref="UniProt/TREMBL:Q9AEM2"
/translation="MNKKLAMAAMAAVTVVGSAAPIFADSTTPGYTVVQNDWKAQVQ
LQGLKXKTIKTSKVSFNGSVGEVTPASSGAKKADRAAEKLYNLVNTQDVKLGDG
DYVDFVYTNLATQITKAEAEAVLTKLQNDKVLINSATDTVKGMVSTQVDSKNV
ANPLKVSMTYIPSAITGSDSGSIAPKTEKTSLLYGTGVDATAGKAITVDTSIN
EAPNGKGVTKYKSPKATQVGGTGTGVLKSDASMAATGTIKRVTSKAEISD
VSSYSIAENLAKVFPKPEVSEAYNAIVALQNDGIESDLVQVNGKYQVTFYPEG
KRLTKSDAIADADSPAKTIIKANKLKDLYVDDLKTNNYTSNVVTVAGDRIET
ATELSKYNSDDKNATTDADVNNIVLVGSTIVDGLVASPLASEKTAPLLTSKDKL
DSVSKSEIKRVNMLKSDGTINTSKKYLKAGGVNSISKOVENELKNMGLKVTLSGEDR
YVETSLAIEGLDNDKAFVGGTGLADAMSIAPVASQLKDGATPIVVVDGKAEIS
DDAKSFLGTSVDVITIGKNSVSKIEISDSATGKTPDRISGDDRQATNAEVLKDDY
FKDGEVYVTVFAKDGSTKEDQLDALAAAPAGRKESPAPIILATDTLSSDQNVAVS
KAVPKDGGTNLVQVGKGIASSVINKMKLLDM"
1. 72
/gene="alp"
/evidence=experimental
73. 1023
/gene="alp"
/product="lower S-layer protein"
/evidence=experimental
1024. 2142
/gene="alp"

/product="upper S-layer protein" /evidence=experimental		Query Match Best Local Similarity 62.4%; Score 483.2; DB 1; Length 2145; Matches 911; Conservative 0; Mismatches 433; Indels 108; Gaps 6;			
QY	394	ACTGTTGAAATTAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTGTGTATCTAGTAGTCG 453	1705	GCTAAGAGTTTCTTAGGAACCTTCGATGTTGATATAATAGTGGAAAAATAGCGTATCT 1764	
Db	784	ACTATAAAAGTTAGAGTTACAAAGTGCAAAAGAGAAATCTATTGATGTGGATTCAAGTTCA 843	1408	GAAAGATGGAAGAGCTTATATCAGATGCTACTGCTAAAGCGGTTTACAAGAGTTAAAGGC 1467	
QY	454	AAAAATAGTGCAAAAGATATAGCTGAAAATATGTGTTGAAGACAAAGACTTAGAAAAAT 513	1765	AAAGAGATTGAAGAGTCAATAGATAGTGCACACTGGAAAAAATCCAGATAGATAAGTGA 1824	
Db	844	TATATTAGTGTCTGAAATTTAGCTTAAAAATATGTTTAAATCCCTAAAGAGGTTTCTGAA 903	1468	GACGATAGACAGACACTTAACCTCTGAAGTTTATAAAAAACATATTATGCTTAATGATACTGAA 1527	
QY	514	GCATTAATAAATCTATAAATCCCTCAGATTTTCAGTAAACCTGATGTTACTATCAAGTAGTT 573	1825	GATGACAGACAGCAACTTAATGCTGAAGTTTTTAAAA----- 1860	
Db	904	GCTTATAATGCAATAGTTGCAATTCACAAATGATGGAATAGAAATCTGATTTAGTACAATTA 963	1528	ATAGCTAAAGCTGCAGTTTTTAGATAAAGATTCAGGTGCTTCAAGTAGTGATGCAGAGTA 1587	
QY	574	CTTTATCCAAAAGGAAGAGATTACAAGTTTTCTCAACTTTATAG---AGCTACAAATTA 630	1861	-----GAAGATGATTAATTTTCAAGATGGTGAAGTT 1890	
Db	964	GTTAATGGAAATATATCAAGTTATTTCTATCCAGAAGGAAAAAGATTAGAACTAAATCT 1023	1588	TTTAAATTTCTATAGCTAAAGATGGATCTACAAAGAAAGATCAATTTAGTTGATGCAATTA 1647	
QY	631	AATGAAGGAACCTGCATATGTTAATACACAGTAATTAATTAACCTTAAAAATCTACTAGTAAG 690	1891	GTGAATTAATCTTTGTTGCAAAAGATGGTTCTACTTAAAGAAAGATCAATTTAGTAGTCAATTA 1950	
Db	1024	GCAGATATAATAGCTGATCGATATAGTCCAGCTTAAATTAACATATAAAGCTAATAAATTA 1083	1648	GCAGTAGGAGCTGTGCTGG-----ATATAAACTTGTCTCCAGTTGTATTAGCTACT 1698	
QY	691	AGTAATTTAAAGACTGCAGTAGAAGAGTTTCAAAAGTTTGAATGCTAGTTATTCTTAATCT 750	1951	GCAGCAGCACCAATAGCAGGTAGATTTAAGGAGTCTCCAGCTCCAATCATACTAGCTACT 2010	
Db	1084	AAAGATTAAAGATTTATGATGATGATTTAAACAATCAATTAATTAATTAATGTT 1143	1699	GATTCCTTTATCTTCGATCAATCGGTTGCTATTAAGCAAAAGTTGTAGGAGAAAAATATTCT 1758	
QY	751	ACAACCTTTAGCTGGTATGACAGAGATACAAACAGCTATATAGATAGTAAGTAAGAAATTAAC 810	2011	GATACCTTTATCTTCGACCAAAATGTAGCTGTAGTAAAGCAGTTCTCTTAAGATGTGGA 2070	
Db	1144	GTAACAGTAGCAGGAGAGATAGATAGAACTGCTATAGAAATTAAGTAGTAATAATTA 1203	1759	AAAGATTTAAACACAAAGTTGGTCAAGGAATAGCTAATTCAGTTATAACAAAAATGAAGAT 1818	
QY	811	ATAATGATGGCGAGAAATCAGATCAATCAGCTGATGTTTAAAGAGAAATGTTTAAAAATGTT 870	2071	ACTAACTTAGTTCAAGTAGGTAAAGGTATAGCTTCTTCAGTTATAACAAAAATGAAGAT 2130	
Db	1204	AATTCGTGATGATAAAATGCAATAACTGATGATGCAAGTTAT-----AATATA 1251	1819	TTATTAGATATG 1830	
QY	871	GTATTAGTAGTGGCAAACTGCACTAGTAGATGGAATGATTGCGGCTCCTTTAGCAGCAAA 930	2131	TTATTAGATATG 2142	
Db	1252	GTATTAGTTGGATCTACATCTATAGTTGATGCTGTTGTCATCACCATTAGCTTCAGAA 1311	RESULT 11		
QY	931	AAAGATGCTCCATATTATTAATCTCAAAAGATAAATTAAGATTCGTCAGTAAAACTGAA 990	AX259456	Sequence 4 from Patent WO0173040.	
Db	1312	AAAAAGCTCCATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1371	LOCUS	AX259456	
QY	991	ATAAGAGAGTTTATAGACTTAAAACTTCAACAGAGTA--ACAGGAAAAACGTTTAT 1047	DEFINITION	Sequence 4 from Patent WO0173040.	
Db	1372	ATAAAAGAGTTATGAACCTTAAAGAGTGTATCTGTTATTAATTAATCTTCTTAAAAAGTTTAT 1431	ACCESSION	AX259456	
QY	1048	ATAGCTGTGGAGTTAATAGTGTATCTAAAGAGTTGTTAAACAGAAATTAAGAAATCAATGGGA 1107	VERSION	AX259456.1 GI:16508624	
Db	1432	TTAGCTGGTGGAGTTAATTTCTATATCTAAAGATGTAGAAATGAATTTGAAATATGGGC 1491	KEYWORDS	Clostridium difficile	
QY	1108	TTAAAGTTGAAAGATTCAGCTGATGATAGATAGAACTTCTTTAAAAATAGCAGGT 1167	SOURCE	Clostridium difficile	
Db	1492	CTTAAAGTTTACTAGATTATCAGGAGAAGACAGATACGAACTTCTTTAGCAATAGCTGAT 1551	ORGANISM	Clostridium difficile	
QY	1168	GAAATAGGCTTATGATAAGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1227	REFERENCE	1	
Db	1552	GAAATAGGCTTATGATAAGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1611	AUTHORS	Fairweather, N.P. and Calabi, E.D.	
QY	1228	ATGAGTATAGCTTTCAGTTCTTCTACTAAATTTAGATGTTGTTGTTGTTGTTGTTGTTGTT 1287	TITLE	Clostridium difficile polypeptides and uses thereof	
Db	1612	ATGAGTATAGCTTTCAGTTCTTCT-----CAACTTAAA 1644	JOURNAL	Patent: WO 0173040-A 4 04-OCT-2001;	
QY	1288	AATGGACATGCTACTCCAAATATGTTGTTGATGGAAGCTGATAAATAATCTCGATGAC 1347	IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY & MEDICINE (GB)		
Db	1645	GATGGAGATGCTACTCCAAATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1704	Location/Qualifiers	1. 2145	
QY	1348	TTAGATAGTTTCTTAGGAGCGCTGATGTTAGATATAAATAGTGGATTTCGAGTGTATCT 1407	/organism="Clostridium difficile"		
			/mol_type="unassigned DNA"		
			/db_xref="taxon:1496"		
ORIGIN		Query Match 26.4%; Score 483.2; DB 6; Length 2145; Best Local Similarity 62.7%; Pred. No. 1e-56; Matches 911; Conservative 0; Mismatches 433; Indels 108; Gaps 6;			
QY	394	ACTGTTGAAATTAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTGTGTATCTAGTAGTCG 453	394	ACTGTTGAAATTAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTGTGTATCTAGTAGTCG 453	
Db	784	ACTATAAAAGTTAGAGTTACAAAGTGCAAAAGAGAAATCTATTGATGTGGATTCAAGTTCA 843	Db	784	ACTATAAAAGTTAGAGTTACAAAGTGCAAAAGAGAAATCTATTGATGTGGATTCAAGTTCA 843
QY	454	AAAAATAGTGCAAAAGATATAGCTGAAAATATGTGTTGAAGACAAAGACTTAGAAAAAT 513	QY	454	AAAAATAGTGCAAAAGATATAGCTGAAAATATGTGTTGAAGACAAAGACTTAGAAAAAT 513
Db	844	TATATTAGTGTCTGAAATTTAGCTTAAAAATATGTTTAAATCCCTAAAGAGGTTTCTGAA 903	Db	844	TATATTAGTGTCTGAAATTTAGCTTAAAAATATGTTTAAATCCCTAAAGAGGTTTCTGAA 903
QY	514	GCATTAATAAATCTATAAATCCCTCAGATTTTCAGTAAACCTGATGTTACTATCAAGTAGTT 573	QY	514	GCACCTAAAAAATATAAATGCTCAGATTTTCAGTAAAACTGATGTTACTATCAAGTAGTT 573

Db	904	GCCTTATAATGCAATAGTGTGCATTACAAATATGATGGAATAGAAATCTGATTTAGTACAAATTA	963
Qy	574	CTTTATCCAAAAGAAAGAGATTACAAGGTTTCTCAACTTATAG---AGCTACAAATTTAT	630
Db	964	GTTAAATGGAAAATATCAAGATTATTTTCTATCCAGAAAGGAAAAGATTAGAAAATAAACTT	1023
Qy	631	AATGAAGGAACTGCATATGTTATATACACCAAGTAAATTAATTAATCTCAAAATCTACTAGTAAG	690
Db	1024	GCAGATATATAGCTGATGAGATAGTCCAGCTAAATTAATTAATTAAGCTAATAATTA	1083
Qy	691	AGTAATTTAAAGACTGCGAGTAGAAGATTACAAAATTTGAATGCTAGTATTCTTAATCT	750
Db	1084	AAAGATTTAAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1143
Qy	751	ACAACTTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	810
Db	1144	GTAACAGTAGCAGGAGAAGATAGAAATAGAAATCTGCTATAGAAATTTAGTAGTAATTTAT	1203
Qy	811	AATAATGATGGCGAGAAATCAGATCAATTCAGCTGATGTTAAAGAGAATGTTAAATAATGTT	870
Db	1204	AATTCGATGATAAATGCAATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1251
Qy	871	GTATTAGTAGTGCATATGCACTAGTAGATGATGATGATGATGATGATGATGATGATGATGATGAT	930
Db	1252	GTATTAGTTGGATCTACATCTATAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT	1311
Qy	931	AAAGATGCTCCACTATTTAATTAATCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT	990
Db	1312	AAAAAGCTCCATTTATTTAATTAATCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT	1371
Qy	991	ATAAGAGATTTTATAGATTTTAAATCTTCAAGATTAATTAATTAATTAATTAATTAATTAATTAAT	1047
Db	1372	ATAAAAGAGTTTATGAATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1431
Qy	1048	ATAGCTGCTGGAGTTTAAATAGTATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGAT	1107
Db	1432	TTAGCTGCTGGAGTTTAAATAGTATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGAT	1491
Qy	1108	TTAAAGTTGAAAGATTTCTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1167
Db	1492	CTTAAAGTTTACTAGATTTATCAGGAGAAAGACAGATACGAAATCTTCTTTAGCAATAGCTGAT	1551
Qy	1168	GAAATAGGCTTATGATATGATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1227
Db	1552	GAAATAGGCTTATGATATGATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1611
Qy	1228	ATGATATAGCTTCAAGTTGCTTCTATCTAAATTTAGATGTTAGTGTGTTGTAGATAGAACTA	1287
Db	1612	ATGATATAGCTTCAAGTTGCTTCTATCTAAATTTAGATGTTAGTGTGTTGTAGATAGAACTA	1644
Qy	1288	AATGGACATGCTACTCCAAATAGTTGTTGTAGATGGAAGGTTGATTAATAATCTGATGATGAT	1347
Db	1645	GATGGAGATGCTACTCCAAATAGTTGTTGTAGATGGAAGGTTGATTAATAATCTGATGATGAT	1704
Qy	1348	TTAGATAGTTTCTTAGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1407
Db	1705	GCTAAGAGTTTCTTAGGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1764
Qy	1408	GAAAGATGGAAGAGCTATATCAGATGCTACTGTTAAAGCGTTTACAGAGATTGTAAGGC	1467
Db	1765	AAAGAGATTGAGAGTCAATAGATAGTGCATCTGGAAGGTTTACAGATAGAAATGATGATGAT	1824
Qy	1468	GACGATAGACAGACATTAATCTGGAAGTTTAAATAATTAATTAATTAATTAATTAATTAATTAAT	1527
Db	1825	GATGACAGACAGACATTAATCTGGAAGTTTAAATAATTAATTAATTAATTAATTAATTAATTAAT	1860
Qy	1528	ATAGCTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1587
Db	1861	-----GAAATGATTTATTTTCAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1890
Qy	1588	TTTAAATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1647
Db	1891	GTGAATTTCTTTGTTGGAAGATGTTTCTACTTAAGAGATCAATTTAGTAGTAGTAGTAGTAGTAG	1950
Qy	1648	GCAGTAGGAGCTGTTGCTGG-----ATATAAATCTGCTCCAGTTGTGTATTAGCTACT	1698
Db	1951	GCAGCAGCACCAATAGCAGGTAGATTAAAGGAGTCTCCAGCTCCAATCATCTAGCTACT	2010
Qy	1699	GATCTTTATCTTCTGATCAATCGGTTGCTATTAAGCAAGTTGTAGGAGAAAATATTTCT	1758
Db	2011	GATCTTTATCTTCTGATCAATCGGTTGCTATTAAGCAAGTTGTAGGAGAAAATATTTCT	2070
Qy	1759	AAAGATTTAAACAAGATTGCTCAAGCAATAGCTTAATTCAGTTTATAAACAATAAGAAAT	1818
Db	2071	ACTAATCTAGTTCAAGTAGGTAAAGTATAGCTTCTTCAAGTTATAAACAATAAGAAAT	2130
Qy	1819	TTATTAGATATG 1830	
Db	2131	TTATTAGATATG 2142	
RESULT 12			
AX512907			
LOCUS	AX512907	2145 bp	DNA linear PAT 03-OCT-2002
DEFINITION	Sequence 7 from Patent WO02062379.		
ACCESSION	AX512907		
VERSION	AX512907.1	GI:23504066	
KEYWORDS	Clostridium difficile		
SOURCE	Clostridium difficile		
ORGANISM	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.		
REFERENCE	1		
AUTHORS	Doyle, R., Kelleher, D., Windle, H. J., Walsh, J. B. and Deirdre, N. E.		
TITLE	Clostridium difficile vaccine		
JOURNAL	Patent: WO 02062379-A 7 15-AUG-2002;		
	THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH		
	(IE)		
FEATURES	Location/Qualifiers		
source	1..2145		
	/organism="Clostridium difficile"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:1496"		
ORIGIN			
Query Match	26.2%;	Score 480;	DB 6; Length 2145;
Best Local Similarity	62.6%;	Pred. No. 2.8e-56;	
Matches	909;	Conservative 0;	Mismatches 435; Indels 108; Gaps 6;
Qy	394	ACTGTTGAAATTAAGTAGTAGCAGCTAGTGTGAAAAACAGTAGTTGTATCTAGTGCG	453
Db	784	ACTATAAAGTTAGAGTTACAGTGCAGTCAAAAGAGAAATCTATTGATGGGATTCAGGTTCA	843
Qy	454	AAAAATAGTGCAAAAGATATAGCTGAAAAATATGTGTTTCAAGACAAAGACTTAGAAAAAT	513
Db	844	TATATTAGTGCTGAAAATTTAGCTAAAAATATGTTTAAATCTCTAAGAGGTTTCTGAA	903
Qy	514	GCACATAAACTATAAATGCTCAGATTTTCAGTAAAACTGATGTTACTATCAAGTAGTT	573
Db	904	GCATTATAATGCAATAGTTGCAATACAAATATGATGGAATAGAACTGATTTAGTACAATTA	963
Qy	574	CTTTATCCAAAAGGAAAGAGATTACAAGGTTTCTCAACTATAG---AGCTACAAATAT	630
Db	964	GTTAAATGGAAAATATCAAGTTATTTTCTATCCAGAAAGAAAAGATTAGAAATCT	1023
Qy	631	AATGAAGGAACTGCATATGTTATATACACCAAGTAAATTAATTAATCTCAAAATCTACTAGTAAG	690
Db	1024	GCAGATATATAGCTGATGAGATAGTCCAGCTAAATTAATTAATTAAGCTAATAATTA	1083
Qy	691	AGTAATTTAAAGACTGCGAGTAGAAGATTACAAAATTTGAATGCTAGTATTCTTAATCT	750
Db	1084	AAAGATTTAAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1143
Qy	751	ACAACTTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	810
Db	1144	GTAACAGTAGCAGGAGAAGATAGAAATAGAAATCTGCTATAGAAATTTAGTAGTAATTTAT	1203

QY 811 AATAATGATGGCGAGAAATCAGATCAATTCAGCTGATGTTAAAGAGAAATGTTAAAAATGTT 870
Db 1204 AATTTCTGATGATAAAAAATGCAATAAATGATGATGCGAGTTAAAT-----AATATA 1251
QY 871 GTATTAGTAGTGCAAAATGCACCTAGTAGATGATAGTTAGTTGCGGCTCTTTTAGCAGCAGAA 930
Db 1252 GTATTAGTTGGATCTACATCATATAGTTGATGGTCTTGTGGCATCACCATTAGCTTCAGAA 1311
QY 931 AAAGATGCTCCACTATTATTAACCTCAAAGATAAATTAGATTGCTCAGTAAATCTGAA 990
Db 1312 AAAACAGCTCCATTAATTAACCTCAAAGATAAATTAGATTCACTAGTAAATCTGAG 1371
QY 991 ATAAGAGAGTTTGAAGCTTAAACCTCAACAGAAAGTA---ACAGGAAAAACAGTTTAT 1047
Db 1372 ATAAAAAGAGTTATGAACCTAAAGAGTGATCTGGTATAAATACCTTCTAAAAAAGTTTAT 1431
QY 1048 ATAGCTGGTGGAGTTAATAGTGTATCTAAAGAGTTGTAACAGAAATTAGAATCAATGGCA 1107
Db 1432 TTAGCTGGTGGAGTTAATTTCTATATCTAAAGATGTAAGAGATGAATTGAAAAATATGGGC 1491
QY 1108 TTAAGAGTTGAAAGATTCTCAGGTGATGATAGATATGAAACTTCTTTAAAAATAGCAGGT 1167
Db 1492 CTTAAGATTACTAGATTATCAGGAGAAGACAGATACGAACTTCTTTAGCAATAGCTGAT 1551
QY 1168 GAAATAGGCTTAGAATATGATAAGGCTTATGATGTTGGTGGAAACAGGATTAGCAGATGCC 1227
Db 1552 GAAATAGGCTTCTGATAATGATAAGCATTGTTAGTTGGTGGTACTGCTGATGGCAGATGCT 1611
QY 1228 ATGAGTATAGCTTCAGTCTCTCTACTAAATTAGATGGTAAATGCTGTCTAGATGAACA 1287
Db 1612 ATGAGTATAGCTTCAGTCTCTCT-----CAACTTAAA 1644
QY 1288 AATGGACATGCTACTCCATAGTTGTTGTAGATGGAAAAAGCTGATAAAAATATCTGATGAC 1347
Db 1645 GATGGAGATGCTACTCCATAGTTAGTTGTAGATGGAAAGCAAAAGAAATAGTGTATGAT 1704
QY 1348 TTAGATAGTTTCTTAGAAGCGCTGATGATAGATATATATAGTTGGATTTCCAAAGTGTATCT 1407
Db 1705 GCTAAGAGTTTCTTAGGAACTTCTGATGTTGTATATATAGTTGGTGGAAAAAATAGCGTATCT 1764
QY 1408 GAAAGATGGAAGCACTATATCAGATGCTACTGTTAAAGGCGTTACAGAGTTTAAAGGC 1467
Db 1765 AAGAGATTGAAGAGTCAATAGATAGTGCACCTGGAAAAAATCCAGATAGAAATAGTGGA 1824
QY 1468 GACGATAGACAAGACACTAACTCTGAAGTTATAAAAAACATATATGCTAATGATACCTGAA 1527
Db 1825 GATGACAGACAAGCAACTAATGCTGAAGTTTAAAA----- 1860
QY 1528 ATAGCTAAAGCTGCAGTTTATAGATAAAGATTGAGTGTCTCAAGTGTATGATGATGAGGTA 1587
Db 1861 -----GAAGATGATTATTTCAAGATGGTGAAGTT 1890
QY 1588 TTTAATTTCTATGCTAGCTAAAGATGATCTACAAAAGAGATCAATTTAGTTGATGATTA 1647
Db 1891 GTGAATTTCTTTGCAAAAGATGGTCTCTACAAAAGAGATCAATTTAGTATGATGATTA 1950
QY 1648 GCAGTAGGAGCTGTTGCTGG-----ATATAAACTTGTCTCAGTTGCTTATTTAGCTACT 1698
Db 1951 GCAGCAGCACCATAAGCAGGTAGATTAAAGAGTCTCCAGCTCCAATCATACTAGCTACT 2010
QY 1699 GATCTTTATCTCTGATCAATCGGTTGCTATTAAGCAAAAGTTGTAGGAGAAAAATTTCT 1758
Db 2011 GATACTTTATCTCTGACCAAAAATGATGCTGTAAGTAAAGCAGTTCTCCTAAAGATGGTGA 2070
QY 1759 AAAGATTTAAACAAGTGTCTCAAGGAATAGCTTAATTTAGTTTATAAACAATAAAGAT 1818
Db 2071 ACTAACTTAGTTCAGAGTAGTAAGGTATAGCTTCTTCAGTTATATAAACAATAAAGAT 2130
QY 1819 TTATTAGATATG 1830
Db 2131 TTATTAGATATG 2142

RESULT 13
AX512901
LOCUS
DEFINITION
Sequence 1 from Patent WO02062379.
ACCESSION
AX512901
VERSION
AX512901.1 GI:23504060
KEYWORDS
Clostridium difficile
SOURCE
Clostridium difficile
ORGANISM
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE
1
AUTHORS
Doyle, R., Kelleher, D., Windle, H.J., Walsh, J.B. and Deirdre, N.E.
TITLE
Clostridium difficile vaccine
JOURNAL
Patent: WO 02062379-A 1 15-AUG-2002;
THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH
(IE)
FEATURES
Location/Qualifiers
source
1..2157
/organism="Clostridium difficile"
/mol_type="unassigned DNA"
/db_xref="taxon:1496"
ORIGIN
Query Match 25.8%; Score 471.8; DB 6; Length 2157;
Best Local Similarity 61.5%; Pred. No. 3.6e-55;
Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;
QY 296 CTCACGCTAAATGGAATGAAGATTATGTAAGACAACTTTAAACAACTTTAGATGCGAGAG 355
Db 728 CAACACCTAGTCTGCTAGCTGTAAGTGGTTTGTAACTAAAGATGATCTGATTTAGCNA 787
QY 356 AATATGCTATTATAGATTTAACTTATAATAATGCTTAAACCTGTTGAAATTTAAAGTAGTAG 415
Db 788 AATCAGGTACTATAAATGTAAGAGTTATAAATGCAAAAGAGAAATCAATTTGATATAGATG 847
QY 416 CAGCTAGTGAAGAAACAGTAGTTGTATCTAGTGATGCGAAAAATAGTGCAGAAAGATATAG 475
Db 848 CAAGCTCATATACATCAGCTG-----AAAATTTAGCTAAAAAGATATGTTTGTGTCAG 901
QY 476 CTGAAAAATATGTTGTTGAAGCAAGAGACTTAGAAAAATGCATAAAACATATAAATGCTCT 535
Db 902 ATGAAATTTCTGAACCATATAAGGCATATAGGCATTAACAATGATGGTATAGATCTA 961
QY 536 CAGATTTTCAAGTAAAACTGATGTTACTATCAAGTAGTTCTTTTATCCAAAAAGAAAGAGAT 595
Db 962 ATTTAGTTTCAGTTAGTTAATGGAATATCAAGTGATTTTATCCAGAGGCTAAAAAGAT 1021
QY 596 TACAAGGTTTCTCAACTTATAGAGCTACAATATTAATGAAGAACTGCATATGTTAATA 655
Db 1022 TAGAACTAAATCAGCA-----AATGATACATAGCTAGTCAAGATA 1063
QY 656 CACCAAGTAAATTAACCTCTAAATCTTACTAGTAGAGTAATTTAAAGACTGCGAGTAGAAG 715
Db 1064 CACCAAGCTTAAGTAGTTATAAAGCTAATAAATTAAGAGTTTAAAGATTTATGATG 1123
QY 716 AGTTCAAAAAATTTGAATGCTAGTTATTCTTAATACTACAACCTTTAGCTGGTGAATGACAGAA 775
Db 1124 ATTTAAAAACATATAATAATACTTATTCAAAATGTTGTAACAGTAGCAGGAGAGATAGAA 1183
QY 776 TACAACAGCTATAGAGATAAGTAAGAAATATTAACAATATGATGGCGAGAAATCAGATC 835
Db 1184 TAGAAACTGCTATAGAAATTAAGTAGTAATAATATATAATTTCTGAT-----GATA 1231
QY 836 ATTCAGCTGATGTTTAAAGAGAAATGTTAAAAATGTTGTTATTAGTAGTGCATAATGCACTAG 895
Db 1232 AAAATGCAATAACTGTAAGCAGTTATGATATAGTATTAGTTGGATCTACATCTATAG 1291
QY 896 TAGATGGATTAGTTGCGGCTCCTTTAGCAGCGAAAAAGATGCTCCACTATTATTAACTTT 955
Db 1292 TTGATGGTCTTGTGTCATCACCATTAGCTTTCAGAAAAAACAGCTCCATTTATTATTAACTTT 1351

QY 956 CAAAGAGTAATAGATTCGTGCTAGTAAATCTGAAATAAAGAGAGTGTTCAGACTTAAAGAA 1015
DB 1352 CAAAGAGTAATAGATTCGTGCTAGTAAATCTGAAATAAAGAGAGTGTTCAGACTTAAAGAA 1411
QY 1016 CTTCAACAGAGTA---ACAGGAAACAGTGTATATATAGCTGGTGGAGTAAATAGTGTAT 1072
DB 1412 GTGACACTGGTATAAATACCTCTAAAGAGTGTATATAGCTGGTGGAGTAAATCTATAT 1471
QY 1073 CTAAGAGTGTGTAACAGATTAAGATCAATGCGATTAAGAGTGTAAAGTGTTCAGCTG 1132
DB 1472 CTAAGAGTGTGTAAGATTAAGATCAATGCGATTAAGAGTGTAAAGTGTTCAGATTAACAGG 1531
QY 1133 ATGATAGATTAAGAACTCTTTAAAGATAGCAGTGAATAGCTTAGATTAATAGTAAGG 1192
DB 1532 AAGACAGATACGAACTCTTTAGCAATAGCTGATGAATAGCTTGTGATTAATAGTAAG 1591
QY 1193 CTTATGATGTTGGTGAACAGGATTAAGCAATGCGATTAAGTGTTCAGTGTTCCTTCTA 1252
DB 1592 CATTTGATGTTGGTGGTACTGGATTAAGCAATGCGATTAAGTGTTCAGTGTTCCTTCT- 1650
QY 1253 CTAATATAGATGTTAGTGTGTTAGATAGAACTGATGATGATGATGATGATGATGATG 1312
DB 1651 -----CAACTTAAAGATGCGATGCTACTCCAAATAGTAG 1684
QY 1313 TTGTAGATGAAGAGCTGATAAATATCTGATGATGATGATGATGATGATGATGATGATG 1372
DB 1685 TTGTAGATGAAGAGCAAGAAATAGTGTGATGATGATGATGATGATGATGATGATGATG 1744
QY 1373 ATGTAGATTAATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1432
DB 1745 ATGTGATTAATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1804
QY 1433 ATGCTACTGTAAGAGGCTTCAAGAGTGTAAAGGCGAGTATGATGATGATGATGATGATG 1492
DB 1805 GTGCACTGCAAACTCCAGATAGAAATAGTGTGATGATGATGATGATGATGATGATG 1864
QY 1493 AAGTTATAAACATATATGCTAATGATGATGATGATGATGATGATGATGATGATGATG 1552
DB 1865 AAGTTTAAAGAGATGATTTTC----- 1890
QY 1553 AAGATTGAGTGTCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1612
DB 1891 -----ACAGTGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1930
QY 1613 GATCTACAAAGAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1667
DB 1931 GTTCTCTAAAGAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1990
QY 1668 -----ATATAAATGCTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1723
DB 1991 TTAAGGAGTCTCCAGCTCCATCATGATGATGATGATGATGATGATGATGATGATGATG 2050
QY 1724 TTGCTATAAGCAAAAGTGTAGGAGAAATATCTTAAAGATTTTAAACAAAGTGTGCTCAAG 1783
DB 2051 TAGCTGTAAGTAAGAGAGTCTTAAAGTGTGAGTAACTTATGATGATGATGATGATGATG 2110
QY 1784 GAATAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1830
DB 2111 GTATAGCTTCTTCAAGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATG 2157

RESULT 14
AX512903 2158 bp DNA linear PAT 03-OCT-2002
LOCUS
DEFINITION Sequence 3 from Patent WO02062379.
ACCESSION AX512903
VERSION AX512903.1 GI:23504062
KEYWORDS
SOURCE Clostridium difficile
ORGANISM Clostridium difficile
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE
1

AUTHORS Doyle, R., Kelleher, D., Windle, H.J., Walsh, J.B. and Deirdre, N.E.
TITLE Clostridium difficile vaccine
JOURNAL Patent: WO 02062379-A 3 15-AUG-2002;
THE COLLEGE OF THE HOLY AND UNDIVIDED TRINITY OF QUEEN ELIZABETH
(IE)
FEATURES Location/Qualifiers
source 1..2158
/organism="Clostridium difficile"
/mol_type="unassigned DNA"
/db_xref="taxon:1496"
ORIGIN
Query Match 25.8%; Score 471.8; DB 6; Length 2158;
Best Local Similarity 61.5%; Pred. No. 3.6e-55;
Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;
QY 296 CTACAGCTAATGGAATGAGATTAATGTAAGACAACTTTAAAGAACTTTAGATGACAGGAG 355
DB 728 CAACACCTAGTCTGTAGCTGTAAGTGGTGTGTAAGATGATGATGATGATGATGATGATG 787
QY 356 AATATGCTATTATAGATTTAACTTATAATAATGCTAAAACTGTGAAATTTAAAGTAGTAG 415
DB 788 AATCAGTACTATTAATGTAAGTGTATTAATGCAAAAGAGAAATCAATTTGATATAGATG 847
QY 416 CAGCTAGTGAAGAAACAGATGTTGTAATCTAGTATGCGAAATAATAGTGCAGAAAGATATAG 475
DB 848 CAAGCTCATATACATCAGCTG-----AAAAATTAGCTAAAGATATGTTATTTGATCCAG 901
QY 476 CTGAAAAATATGTTGTTGAAGCAAGACTTAGAAAAATGCACCTAAAACTATATAATGCCT 535
DB 902 ATGAATTTCTGAAGCATATAAGCAATAGTAGTACATTAACAATGATGGTATAGATGCTTA 961
QY 536 CAGATTTTCAGTAAACTGATAGTACTATCAAGTAGTCTTTATCCAAAGGAGAAAGAGAT 595
DB 962 ACTTAGTTCAGTGTATTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAG 1021
QY 596 TACAAGGTTTCTCAACTTATAGAGCTACAAATTTAATGAAGAACTGCAATATGTTAATA 655
DB 1022 TAGAACTAAATCAGCA-----AATGATACAATAGCTAGTCAAGATA 1063
QY 656 CACAGTAAATTAATCTTAATCTACTAGTAAAGATTAATTTAAAGCTGCGAGTAGAAG 715
DB 1064 CACCAGCTAAGTAGTATAAAGAGCTAAATAATTTAAAGATTTTAAAGATTTATGATGATG 1123
QY 716 AGTTACAAAAATGGAATGCTAGTATTCTAATACTACAACTTTAGCTGGTGTGATGACAGAA 775
DB 1124 ATTTAAAGAACTAATAATATATCTTTTCAATGTTGTACAGTAGCAGGAGAGATAGAA 1183
QY 776 TACAAACAGCTATAGAGATAAGTAAAGATATTACAAATATATGATGGCGAGAAATCAGATC 835
DB 1184 TAGAACTGCTATAGAAATTAAGTAGTAAATATTATAATTTCTGAT-----GATA 1231
QY 836 ATTCAGCTGATGTTAAAGAGAAATGTTAAAGATCTGTTATAGTATAGTGGTCAATGCTAG 895
DB 1232 AAAATGCAATPACTGATTAAGAGCAATTAATGATATAGTATTAGTGGATCTACATCTATAG 1291
QY 896 TAGATGATGATGTTGGGCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTT 955
DB 1292 TTGATGGTCTTGTGCAATCACCATTGCTTCAGAAAAACAGCTCCCATTTATTAACTT 1351
QY 956 CAAAAGATAAATTAGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1015
DB 1352 CAAAAGATAAATTAGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1411
QY 1016 CTTCAACAGAGTA---ACAGGAAACAGTGTATATATAGCTGTTGAGTGTAAATAGTGTAT 1072
DB 1412 GTGACACTGCTGATAAATACTTCTAAAAAGTGTATTTAGCTGGTGGAGTAAATTTCTATAT 1471
QY 1073 CTAAGAGTGTGTAACAGAAATAGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 1132
DB 1472 CTAAGAGTGTGTAAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 1531
QY 1133 ATGATAGATGATGAACCTTCTTTAAAAATAGCAGGTGAATAGCTTAGATTAATGATAAGG 1192


```

Db 1685 TTCTAGATGAAAGCAAAAGAAATAAGTGATGATGCTAAGAGTTTCTTAGGAACTTCTG 1744
Qy 1373 ATCTAGATATAATAGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATATCAG 1432
Db 1745 ATGTTGATATAATAGTGGAAAAAATAGCGTATCTAAAGAGATTGAAGAGTCAATAGATA 1804
Qy 1433 ATGCTACTGGTAAAGCGTTTACAAGAGTTTAAAGCGACGATAGACAAAGACACTAACTCTG 1492
Db 1805 GTGCAACTGGAAAACTCCAGATAGATAAGTGGAGATGATAGCAAGCAACTAATGCTG 1864
Qy 1493 AAGTTATAAAAAACATATATATGCTTAATGATAGTGAATAAGCTGAAATAGCTAAAGCTGCAGTTTTAGATA 1552
Db 1865 AAGTTTTAAAAAGAGATGATTATTC----- 1890
Qy 1553 AAGATTCAGGTGCTTCAAGTAGTGTGATGCAGGAGTATTTAATTTCTATGTAGCTAAAGATG 1612
Db 1891 -----ACAGATGGTGAAGTTGTGAATTTACTTTTGTGCAAAAGATG 1930
Qy 1613 GATCTACAAAAGAGATCAATTAGTTGATGCATTTAGCAGTAGGAGCTGTTGCTGG----- 1667
Db 1931 GTTCTACTAAGAGAGATCAATTTAGTAGTGCCTTAGCAGCAGCACCACCAATAGCAGGTAGAT 1990
Qy 1668 ----ATATAAACTTGCTCCAGTTGTATTAGTACTGATTTCTTATCTTCTGATCAATCGG 1723
Db 1991 TTAAGGAGTCTCCAGCTCCAATCATACTAGTACTGATACTTTTATCTTTCTGACCAAAATG 2050
Qy 1724 TTGCTATAAGCAAGTTGTAGGAGAAAAATATTCTAAAGATTTAAACACAGTTGGTCAAG 1783
Db 2051 TAGCTGTAAAGTAAAGCAGTTTCCTAAAGATGGTGGAACTTAAGTTTCAAGTAGGTAAAG 2110
Qy 1784 GAATAGCTAATTCAGTTTATAAAACAAAATGAAAGATTTTATTAGATATG 1830
Db 2111 GTATAGCTTCTTCAGTTATATAACAAAATGAAAGATTTTATTAGATATG 2157

```

Search completed: October 6, 2005, 08:36:58
Job time : 8008 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2005, 04:50:47 ; Search time 1014 Seconds
(without alignments)
10683.557 Million cell updates/sec

Title: US-10-068-870A-4
Perfect score: 1830
Sequence: 1 atgaaaaaaagaatttagc.....tgaaagattattagatatg 1830

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1980s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1830	100.0	1830	6	ABs65376 Clostridi
2	1830	100.0	1830	6	ABs65368 Clostridi
3	542	29.6	2268	6	ABs65367 Clostridi
4	535.6	29.3	2271	6	ABs65378 Clostridi
5	503.4	27.5	2217	4	AAI65841 Nucleotid
6	503.4	27.5	2217	6	ABs65372 Clostridi
7	503.4	27.5	2217	6	ABs65380 Clostridi
8	483.2	26.4	2145	4	AAI65839 Nucleotid
9	480	26.2	2145	6	ABs65371 Clostridi
10	480	26.2	2145	6	ABs65381 Clostridi
11	471.8	25.8	2157	6	ABs65369 Clostridi
12	471.8	25.8	2157	6	ABs65374 Clostridi
13	471.8	25.8	2157	6	ABs65373 Clostridi
14	471.8	25.8	2157	6	ABs65375 Clostridi
15	471.8	25.8	2158	6	ABs65377 Clostridi
16	471.8	25.8	2158	6	ABs65382 Clostridi
17	471.8	25.8	2160	4	AAI65840 Nucleotid
18	471.8	25.8	48551	6	AAI65840 Nucleotid
19	470.2	25.7	2157	6	ABs65370 Clostridi
20	470.2	25.7	2158	6	ABs65379 Clostridi

21	145.2	7.9	1893	8	ACA28989	Prokaryot
22	112.4	6.1	3399	2	AA705868	Chicken 1
23	104	5.7	1503	8	ACA28912	Prokaryot
24	99.6	5.4	1830	8	ACA28975	Prokaryot
25	95.8	5.2	1686	2	AAQ87587	DNA encod
26	95.6	5.2	5361	2	AA778868	P. falcip
27	95.6	5.2	5529	6	ABK50883	DNA encod
28	95.6	5.2	6152	2	AA778867	P. falcip
29	92	5.0	4997	8	ACA47636	Prokaryot
30	90.2	4.9	3095	2	AAQ03875	Sequence
31	90.2	4.9	3579	3	AA700099	Plasmodi
32	88.4	4.8	5253	8	ACA29579	Prokaryot
33	87.6	4.8	7442	4	AA846686	Tumour su
34	86.2	4.7	7446	8	ACA19708	Prokaryot
35	86.2	4.7	7446	10	ACF05848	Methicill
36	85.2	4.7	4965	8	ADA89806	Staphyloc
37	85.2	4.7	7434	4	AA552179	Staphyloc
38	85.2	4.7	7434	8	ABT14965	Pathogen
39	85.2	4.7	7437	8	ACF73459	Staphyloc
40	85.2	4.7	7437	4	AA555232	Staphyloc
41	85.2	4.7	8155	2	AAV74374	Staphyloc
42	84.8	4.6	1300	12	ADP85917	Synthetic
43	84.2	4.6	4911	8	ABT15076	Pathogen
44	82.2	4.5	2370	4	AAH28330	Nucleotid
45	81.8	4.5	6292	4	AA846735	Tumour su

ALIGNMENTS

RESULT 1

ABs65376

ID ABs65376 standard; DNA; 1830 BP.

AC ABs65376;

DT 15-NOV-2002 (first entry)

DE Clostridium difficile strain 172450 slpa DNA sequence.

XX Vaccine; Clostridium difficile; immunogenic; immunoreactivity;

KW C. difficile infection; immune response; C. difficile associated disease;

KW passive immunotherapy; passive vaccination; interleukin; adjuvant;

KW C. difficile vaccine; antibacterial; slpa; ds.

XX Clostridium difficile.

XX WO200262379-A2.

XX 15-AUG-2002.

XX 11-FEB-2002; 2002WO-1E000017.

XX 09-FEB-2001; 2001IE-00000137.

XX (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.

XX Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;

XX WPI; 2002-643389/69.

XX Novel vaccine useful for treatment/prophylaxis of Clostridium difficile

XX associated diseases, comprises Clostridium difficile Genes or

XX peptide/polypeptides or its derivative, fragment, mutant or variant.

XX Claim 12; Page 70-71; 85pp; English.

XX The present invention relates to a new vaccine for treatment/prophylaxis

XX of Clostridium difficile associated disease. The vaccine of the invention

XX comprises C. difficile gene or peptide/polypeptide or derivative,

XX fragment, mutant or variant of the peptide/polypeptide which is

XX immunogenic in humans or to which immunoreactivity is detected in

XX individuals who have recovered from C. difficile infection. The molecules

XX CC

CC of the invention are used in prophylaxis or treatment of C. difficile
 CC associated disease, where the medicament is a vaccine. The invention can
 CC be used by administering the vaccine preparation to a host to raise an
 CC immune response. The invention is further useful in the preparation of a
 CC medicament for treatment or prophylaxis of C. difficile infection or C.
 CC difficile associated disease. The invention is also useful in passive
 CC immunotherapy for established C. difficile infection, for the eradication
 CC of C. difficile associated disease, or for passive vaccination of an
 CC individual with C. difficile infection. Interleukin is useful as an
 CC adjuvant in C. difficile vaccine. The present nucleic acid sequence
 CC represents a Clostridium difficile slpA DNA sequence of the invention
 XX
 SQ Sequence 1830 BP; 716 A; 220 C; 362 G; 532 T; 0 U; 0 Other;

Query Match		100.0%;	Score 1830;	DB 6;	Length 1830;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1830;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAAAAAAGAAATTTAGCAATGGCTATGGCAGCTGTTACTGTAGTAGTGTCTGCTGCT	60		
DB	1	ATGAAAAAAGAAATTTAGCAATGGCTATGGCAGCTGTTACTGTAGTAGTGTCTGCTGCT	60		
QY	61	CCAGTTTTTTCGAGCAGCTTTCAGATGTAATATCACTACAAGATGGTACAAATGATAAGTAT	120		
DB	61	CCAGTTTTTTCGAGCAGCTTTCAGATGTAATATCACTACAAGATGGTACAAATGATAAGTAT	120		
QY	121	ACAGTATCAAACTAAGCTAGTGAATAGTAAGGATATTTTAGCAGCACAAAATTTA	180		
DB	121	ACAGTATCAAACTAAGCTAGTGAATAGTAAGGATATTTTAGCAGCACAAAATTTA	180		
QY	181	ACAAAGGTCAGTATTTTGAACAAAGATACAAAGTTTACTTCTATGATGCAATGAG	240		
DB	181	ACAAAGGTCAGTATTTTGAACAAAGATACAAAGTTTACTTCTATGATGCAATGAG	240		
QY	241	AAAGATTTCTCAACTCGAGATAAAAAAGTTTATTCAGAACAACTTTTAACTACA	300		
DB	241	AAAGATTTCTCAACTCGAGATAAAAAAGTTTATTCAGAACAACTTTTAACTACA	300		
QY	301	GCTAATGGAATGAAGATTTATGTAAGACAACTTTTAAAGATTTAGATGACGAGATAT	360		
DB	301	GCTAATGGAATGAAGATTTATGTAAGACAACTTTTAAAGATTTAGATGACGAGATAT	360		
QY	361	GCTATTATAGATTTAACTTATATATGCTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT	420		
DB	361	GCTATTATAGATTTAACTTATATATGCTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT	420		
QY	421	AGTGAAGAAACAGTAGTGTATCTAGTATGCGAAAAATAGTGCAAAAAGATATAGCTGAA	480		
DB	421	AGTGAAGAAACAGTAGTGTATCTAGTATGCGAAAAATAGTGCAAAAAGATATAGCTGAA	480		
QY	481	AAATATGTTTTGAAGACAAAGCTTAGAATATGCACTTAAAGCTATTAATGCTCAGAT	540		
DB	481	AAATATGTTTTGAAGACAAAGCTTAGAATATGCACTTAAAGCTATTAATGCTCAGAT	540		
QY	541	TTTCAGTAAACCTGATAGTATCTCAAGTAGTCTTTATCCAAAAGGAAGATATCAAA	600		
DB	541	TTTCAGTAAACCTGATAGTATCTCAAGTAGTCTTTATCCAAAAGGAAGATATCAAA	600		
QY	601	GTTTTCTCAACTATAGAGCTACAAATATATGAAGGAACTGCATATGGTAATACACA	660		
DB	601	GTTTTCTCAACTATAGAGCTACAAATATATGAAGGAACTGCATATGGTAATACACA	660		
QY	661	GTAATATTAACCTCTAAATCTAGTAAGAGTAATTTAAAGCTGCGAGTAGAAGTTA	720		
DB	661	GTAATATTAACCTCTAAATCTAGTAAGAGTAATTTAAAGCTGCGAGTAGAAGTTA	720		
QY	721	CAAAAATTTGAATGCTAGTATTTCTAATCTACAACTTTAGCTGGTGATGACAGATACAA	780		
DB	721	CAAAAATTTGAATGCTAGTATTTCTAATCTACAACTTTAGCTGGTGATGACAGATACAA	780		
QY	781	ACAGCTATAGAGATAAGTAAAGAAATTTACAAATATGATGGCGAGAAATCAGATCATTTCA	840		
DB	781	ACAGCTATAGAGATAAGTAAAGAAATTTACAAATATGATGGCGAGAAATCAGATCATTTCA	840		

RESULT 2
 ABS65368
 ID ABS65368 standard; DNA; 1830 BP.

QY	841	GCTGATGTTAAAGAGAAATGTTTAAAAATGTTGTTATTTAGTAGTGGCAAAATGCATAGTAGAT	900
DB	841	GCTGATGTTTAAAGAGAAATGTTTAAAAATGTTGTTATTTAGTAGTGGCAAAATGCATAGTAGAT	900
QY	901	GGATTTAGTTGCGGCTCCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTCARAA	960
DB	901	GGATTTAGTTGCGGCTCCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTTCARAA	960
QY	961	GATAAATTTAGATTCGTCAGTAAAAATCTGAAAAATAAGAGAGATTTTAGACTTTAAAACTTCA	1020
DB	961	GATAAATTTAGATTCGTCAGTAAAAATCTGAAAAATAAGAGAGATTTTAGACTTTAAAACTTCA	1020
QY	1021	ACAGAAAGTAAACAGGAAAAACAGTTTATATAGCTGCTGGAGTTTAAATAGTGTATCTAAAGAA	1080
DB	1021	ACAGAAAGTAAACAGGAAAAACAGTTTATATAGCTGCTGGAGTTTAAATAGTGTATCTAAAGAA	1080
QY	1081	GTTGTAAACAGAAATTTAGAAATCAATGGGATTTAAAGTTTGAAGATTTCTCAGGTGATGATAGA	1140
DB	1081	GTTGTAAACAGAAATTTAGAAATCAATGGGATTTAAAGTTTGAAGATTTCTCAGGTGATGATAGA	1140
QY	1141	TATGAAACCTTTCTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATATGATAAGGCTTATGTA	1200
DB	1141	TATGAAACCTTTCTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATATGATAAGGCTTATGTA	1200
QY	1201	GTTGTGGAACAGGATTTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTACTAAATTA	1260
DB	1201	GTTGTGGAACAGGATTTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTACTAAATTA	1260
QY	1261	GATGTAATGTTGTTGTAGATAGAACAAATGACATGCTACTCCAAATAGTTTGTGTAGAT	1320
DB	1261	GATGTAATGTTGTTGTAGATAGAACAAATGACATGCTACTCCAAATAGTTTGTGTAGAT	1320
QY	1321	GGAAAGCTGATAAATATCTGATGACTTAGATAGTCTTTCTTAGGAAGCGCTGATGTAGAT	1380
DB	1321	GGAAAGCTGATAAATATCTGATGACTTAGATAGTCTTTCTTAGGAAGCGCTGATGTAGAT	1380
QY	1381	ATAATAGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATATCAGATGCTACT	1440
DB	1381	ATAATAGTGGATTTGCAAGTGTATCTGAAAGATGGAAGAGCTATATATCAGATGCTACT	1440
QY	1441	GTTAAAGCGTTTACAAAGATTTAAAGCGCACATAGACAAAGACATACTCTGAAGTTATA	1500
DB	1441	GTTAAAGCGTTTACAAAGATTTAAAGCGCACATAGACAAAGACATACTCTGAAGTTATA	1500
QY	1501	AAAAATATTATGCTTAATGATATCTGAAATAGCTTAAAGCTGCAAGTTTATAGATAAAGATTCA	1560
DB	1501	AAAAATATTATGCTTAATGATATCTGAAATAGCTTAAAGCTGCAAGTTTATAGATAAAGATTCA	1560
QY	1561	GCTGCTTCAAGTAGTATGAGGAGTATTTAAATTTCTATGCTAGCTTAAAGATGATCTACA	1620
DB	1561	GCTGCTTCAAGTAGTATGAGGAGTATTTAAATTTCTATGCTAGCTTAAAGATGATCTACA	1620
QY	1621	AAAAGAGATCAATAGTTGATGCTATGATGCTAGGAGCTGTTGCTGGATATAAACTGCT	1680
DB	1621	AAAAGAGATCAATAGTTGATGCTATGATGCTAGGAGCTGTTGCTGGATATAAACTGCT	1680
QY	1681	CCAGTTGTTATGCTTACTGATTTCTTTTCTGATCAATCGTTGCTGATATAAGCAAGTT	1740
DB	1681	CCAGTTGTTATGCTTACTGATTTCTTTTCTGATCAATCGTTGCTGATATAAGCAAGTT	1740
QY	1741	GTAGAGAAAAATATTCTTAAAGATTTAAACAGATTTGCTCAAGGAATAGCTTAATTCAGTT	1800
DB	1741	GTAGAGAAAAATATTCTTAAAGATTTAAACAGATTTGCTCAAGGAATAGCTTAATTCAGTT	1800
QY	1801	ATAAACAATAAGAAAGATTTTATAGATATG	1830
DB	1801	ATAAACAATAAGAAAGATTTTATAGATATG	1830

XX AC ABS5368;
 XX DT 15-NOV-2002 (first entry)
 XX DE Clostridium difficile strain 172450 slpA gene.
 XX KW Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
 KW C. difficile infection; immune response; C. difficile associated disease;
 KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
 KW C. difficile vaccine; antibacterial; slpA; gene; ds.
 XX OS Clostridium difficile.
 XX PN WO200262379-A2.
 XX PD 15-AUG-2002.
 XX PP 11-FEB-2002; 2002WO-IE000017.
 XX PR 09-FEB-2001; 2001IE-00000137.
 XX PA (QUEB-) QUEBEN ELIZABETH COLLEGE DUBLIN.
 XX PI Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
 XX DR WPI; 2002-643389/69.
 XX DR P-PSDB; ABG80347.
 XX PT Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
 PT associated diseases, comprises Clostridium difficile genes or
 PT peptide/polypeptides or its derivative, fragment, mutant or variant.
 XX PS Claim 45; Page 29-32; 85pp; English.
 XX The present invention relates to a new vaccine for treatment/prophylaxis
 CC of Clostridium difficile associated disease. The vaccine of the invention
 CC comprises C. difficile gene or peptide/polypeptide or derivative,
 CC fragment, mutant or variant of the peptide/polypeptide which is
 CC immunogenic in humans or to which immunoreactivity is detected in
 CC individuals who have recovered from C. difficile infection. The molecules
 CC of the invention are used in prophylaxis or treatment of C. difficile
 CC associated disease, where the medicament is a vaccine. The invention can
 CC be used by administering the vaccine preparation to a host to raise an
 CC immune response. The invention is further useful in the preparation of a
 CC medicament for treatment or prophylaxis of C. difficile infection or C.
 CC difficile associated disease. The invention is also useful in passive
 CC immunotherapy for established C. difficile infection, for the eradication
 CC of C. difficile associated disease, or for passive vaccination of an
 CC individual with C. difficile infection. Interleukin is useful as an
 CC adjuvant in C. difficile vaccine. The present nucleic acid sequence
 CC encodes a Clostridium difficile slpA protein, as described in the
 CC invention
 XX
 SQ Sequence 1830 BP; 716 A; 220 C; 362 G; 532 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1830; DB 6; Length 1830;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1830; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAAAGAAAGAAATTTAGCAATGGCTATGCGAGCTGTTACTGTAGTAGGTTCTGCTGCT 60
 DB 1 ATGAAAGAAAGAAATTTAGCAATGGCTATGCGAGCTGTTACTGTAGTAGGTTCTGCTGCT 60
 QY 61 CCAGTTTTCGAGCAGCTTCAGATGTAATATCACTACAAGATGTCACAAATGATAAGTAT 120
 DB 61 CCAGTTTTCGAGCAGCTTCAGATGTAATATCACTACAAGATGTCACAAATGATAAGTAT 120
 QY 121 ACAGTATCAAACTACTAAAGCTAGTGAATTAAGGATATTTTAGCAGCACAACCTTA 180
 DB 121 ACAGTATCAAACTACTAAAGCTAGTGAATTAAGGATATTTTAGCAGCACAACCTTA 180
 QY 181 ACAACAGGTGCGAGTTATTTTGACAAAGATACAAAGATTACTTCTATGATGCAATGAG 240

181 ACAACAGGTGCGAGTTATTTTGACAAAGATACAAAGATTACTTCTATGATGCAATGAG 240
 241 AAAGATTCTTCAACTCCAACTGGAGATAAAAAGTTTATTTCAGAACAACTTTTAACCTACA 300
 241 AAAGATTCTTCAACTCCAACTGGAGATAAAAAGTTTATTTCAGAACAACTTTTAACCTACA 300
 301 GCTAATGAAATGAAGATTATGTAAGACAACTTTTAAAAAATTTAGATGCGAGAGAAATAT 360
 301 GCTAATGAAATGAAGATTATGTAAGACAACTTTTAAAAAATTTAGATGCGAGAGAAATAT 360
 361 GCTATTATAGATTAACTTTAATGCTTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT 420
 361 GCTATTATAGATTAACTTTAATGCTTAAACCTGTTGAAATTTAAAGTAGTAGCAGCT 420
 421 AGTGAAGAAAAACAGTAGTTGTATCTAGTGTATGCGAAAAATAGTGCAGAAAGATATAGCTGAA 480
 421 AGTGAAGAAAAACAGTAGTTGTATCTAGTGTATGCGAAAAATAGTGCAGAAAGATATAGCTGAA 480
 481 AAATATGTTGTTGAAGACAAAGACTTAGAAAAATGCACTAAAAAATATATAATGCTCAGAT 540
 481 AAATATGTTGTTGAAGACAAAGACTTAGAAAAATGCACTAAAAAATATATAATGCTCAGAT 540
 541 TTCAGTAAAACTGATAGTTACTATCAAGTAGTTCTTTATCCAAAAGGAAGAGATTACAA 600
 541 TTCAGTAAAACTGATAGTTACTATCAAGTAGTTCTTTATCCAAAAGGAAGAGATTACAA 600
 601 GGTTCCTCAACTTATAGAGCTACAAATATATGAAGGAACTGTCATATGTTAATACACCA 660
 601 GGTTCCTCAACTTATAGAGCTACAAATATATGAAGGAACTGTCATATGTTAATACACCA 660
 661 GTAATATTAACTCTAAAAATCTACTAGTAGAGTAAATTTAAAGACTGTCAGTAGAAGAGTTA 720
 661 GTAATATTAACTCTAAAAATCTACTAGTAGAGTAAATTTAAAGACTGTCAGTAGAAGAGTTA 720
 721 CAAAAATTTGAATGCTAGTTATTCTAATACTACAACTTTAGCTGGTGGTATGACAGAAATACA 780
 721 CAAAAATTTGAATGCTAGTTATTCTAATACTACAACTTTAGCTGGTGGTATGACAGAAATACA 780
 781 ACAGCTATAGAGATAGTAAAGAAATATTAACAATATGATGGGAGAGAAATCAGATCAATTC 840
 781 ACAGCTATAGAGATAGTAAAGAAATATTAACAATATGATGGGAGAGAAATCAGATCAATTC 840
 841 GCTGATGTTAAAGAGAAATCTTAAAAATGTTGTAATAGTAGGTGCAAAATGCATAGTAGAT 900
 841 GCTGATGTTAAAGAGAAATCTTAAAAATGTTGTAATAGTAGGTGCAAAATGCATAGTAGAT 900
 901 GGATTAGTTGGGCTCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTTAACTTCAAAA 960
 901 GGATTAGTTGGGCTCTCTTTAGCAGCAGAAAAAGATGCTCCACTATTTAACTTCAAAA 960
 961 GATTAATTTAGATTGCTCAGTAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCT 1020
 961 GATTAATTTAGATTGCTCAGTAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCT 1020
 1021 ACAGAAAGTAAACAGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGTATCTAAAGAA 1080
 1021 ACAGAAAGTAAACAGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGTATCTAAAGAA 1080
 1081 GTTGTAAACAGAAATGAATCAATGGGATTAAAAAGTTGAAAGATTTCTCAGGTGATGATAGA 1140
 1081 GTTGTAAACAGAAATGAATCAATGGGATTAAAAAGTTGAAAGATTTCTCAGGTGATGATAGA 1140
 1141 TATGAAACTCTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGGCTTATGTA 1200
 1141 TATGAAACTCTTTTAAAAATAGCAGGTGAAATAGGCTTAGATAATGATAAGGCTTATGTA 1200
 1201 GTTGGTGGAAACAGGATTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTACTAAATTA 1260
 1201 GTTGGTGGAAACAGGATTAGCAGATGCCATGAGTATAGCTTTCAGTTGCTTCTACTAAATTA 1260
 1261 GATGGTAAATGGTGTGTAGATAGAAACAAATGGACATGCTACTCCAAATAGTGTGTAGAT 1320

Db 1261 GATGGTAATGGTGTCTAGATAGAACAAATGGACATGCTACTCCAAATAGTTCTGTCTAGAT 1320
 Qy 1321 GGAAAAGCTGATAAATATCTGATGACTTATAGATAGTTCTTAGGAAGCGCTGATGATAT 1380
 Db 1321 GGAAAAGCTGATAAATATCTGATGACTTATAGATAGTTCTTAGGAAGCGCTGATGATAT 1380
 Qy 1381 ATAAATAGGTGGATTGTCAGTGTATCTGAAAAGATGGAAGAGCTATATCATCATGCTACT 1440
 Db 1381 ATAAATAGGTGGATTGTCAGTGTATCTGAAAAGATGGAAGAGCTATATCATCATGCTACT 1440
 Qy 1441 GGTAAAGCGGTTACAGAGCTTAAAGCGGACGATAGACAGACACTAACTCTGAACTTATA 1500
 Db 1441 GGTAAAGCGGTTACAGAGCTTAAAGCGGACGATAGACAGACACTAACTCTGAACTTATA 1500
 Qy 1501 AAAACATATATGCTAAATGATCTGAAATAGCTTAAAGCTGCGAGTTTATAGATAAAGATCA 1560
 Db 1501 AAAACATATATGCTAAATGATCTGAAATAGCTTAAAGCTGCGAGTTTATAGATAAAGATCA 1560
 Qy 1561 GGTGCTTCAAGTAGTATGATGACGAGGATTTTAAATTTCTATGTATGATGATGATGATGAT 1620
 Db 1561 GGTGCTTCAAGTAGTATGATGACGAGGATTTTAAATTTCTATGTATGATGATGATGATGAT 1620
 Qy 1621 AAAGAGATCAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 1621 AAAGAGATCAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 Qy 1681 CAGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Db 1681 CAGTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Qy 1741 GTAGAGAAAAATATTTCTAAAGATTTAAACAAAGTTTGTCTGATATATAAATCTGCT 1800
 Db 1741 GTAGAGAAAAATATTTCTAAAGATTTAAACAAAGTTTGTCTGATATATAAATCTGCT 1800
 Qy 1801 ATAAACAAATGAAAGATTTTATAGATATG 1830
 Db 1801 ATAAACAAATGAAAGATTTTATAGATATG 1830
 RESULT 3
 ABS5367
 ID ABS5367 standard; DNA; 2268 BP.
 XX
 AC ABS5367;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Clostridium difficile strain 171500 slpA gene.
 XX
 KW Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
 KW C. difficile infection; immune response; C. difficile associated disease;
 KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
 KW C. difficile vaccine; antibacterial; slpA; gene; ds.
 XX
 OS Clostridium difficile.
 XX
 PN WO200262379-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 11-FEB-2002; 2002WO-IE000017.
 XX
 PR 09-FEB-2001; 2001IE-00000137.
 XX
 PA (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX
 PI Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
 XX
 DR WPI; 2002-643389/69.
 XX
 PR P-PSDB; ABG80346.
 XX
 PT Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
 associated diseases, comprises Clostridium difficile genes or

PT peptide/polypeptides or its derivative, fragment, mutant or variant.
 XX Claim 44; Page 25-28; 85pp; English.
 PS The present invention relates to a new vaccine for treatment/prophylaxis
 CC of Clostridium difficile associated disease. The vaccine of the invention
 CC comprises C. difficile gene or peptide/polypeptide or derivative,
 CC fragment, mutant or variant of the peptide/polypeptide which is
 CC immunogenic in humans or to which immunoreactivity is detected in
 CC individuals who have recovered from C. difficile infection. The molecules
 CC of the invention are used in prophylaxis or treatment of C. difficile
 CC associated disease, where the medicament is a vaccine. The invention can
 CC be used by administering the vaccine preparation to a host to raise an
 CC immune response. The invention is further useful in the preparation of a
 CC medicament for treatment or prophylaxis of C. difficile infection or C.
 CC difficile associated disease. The invention is also useful in passive
 CC immunotherapy for established C. difficile infection, for the eradication
 CC of C. difficile associated disease, or for passive vaccination of an
 CC individual with C. difficile infection. Interleukin is useful as an
 CC adjuvant in C. difficile vaccine. The present nucleic acid sequence
 CC encodes a Clostridium difficile slpA protein, as described in the
 CC invention
 XX
 SQ Sequence 2268 BP; 893 A; 263 C; 443 G; 669 T; 0 U; 0 Other;
 Query Match 29.6%; Score 542; DB 6; Length 2268;
 Best Local Similarity 59.4%; Pred. No. 4.4e-90;
 Matches 1117; Conservative 0; Mismatches 670; Indels 93; Gaps 8;
 Qy 32 CAGCTGTTACTGTAGTAGTGTCTGCTCCAGTCTTTTGGCAGCAGCTTCAGATGTAATAT 91
 Db 401 CAGCAGATGCTATTAATGCTGGACATCTTCAGCAGATGGTGTGTTACAATACTGCGAG 460
 Qy 92 CACTACAAGATGGTACAAATGATAGTATACAGTATCAAAATCACTAAAGCTAGTGAATTAG 151
 Db 461 CTGCTAGTGGTCTACTGAGACAAATTCAGCAGGAACAACTTGCAATGTGCACTATTT 520
 Qy 152 TAAAGGATATTTTAGCAGCACAAAACCTTAAACAACAGGTGCAGTATTTTGAACAAAGATA 211
 Db 521 TTGACACAGCATATACAGATTCATCTGAACCTGCGGTTAAGATTACTATATAAAGCAGATA 580
 Qy 212 CAAAAGTTTACT---TTCTATGATGCAATGAGAAGATTTCTTCAACTCCAACTGGAGATA 268
 Db 581 TGAATGATCTAAATTTGGTAAAGCAGCTGAGACAACTTATCACTGGGCTTACATTTG 640
 Qy 269 AAAAAGTTTATTCAGAACAACTTTTAACTACAGCTAATGGAATAAGAGATTATGTAAGA 328
 Db 641 AAGATGGGTCTACAGAAAAAATTTGTTAAATTTAGGGACAGTGTATATTATAGATATAACTA 700
 Qy 329 CAACCTTTAAAA---AATTTAGATGACGAGAAATATGCTATTATAGATTTAACTTATAATA 385
 Db 701 AAGCTCTTAAACTTACTGTTCTCTGGAAGTAAAGCAACTGTTAAGTTTGTCTGAAAAA 760
 Qy 386 ATGCTAAAACTGTTG-----AAATTAAGTAGTAGCAGCTAGTGA 427
 Db 761 CACCAAGTGCAGTGTTCACACAGTAAACAACAGCTTAGAATAATAATGCTAAAGAA 820
 Qy 428 AAACAGTAGTTGTATCTAGTATGCGAAAAATATAGTGCAAAAAGATATAGCTGAAAAAATG 487
 Db 821 AAACAATAGATATTGACGCTAGTTCTAGTAAAAACAGCACAGATTTAGCTAAAAAATATG 880
 Qy 488 TGTTTGAAGACAAAGACTTGAANAATGCATTAANAACCTATANAAGCTCCCTCAGATTTCACTA 547
 Db 881 TATTTAATAAAGCTGATTTAAATCTCTTTTATAAGTATTAATAAGGATGAAGCAGATA 940
 Qy 548 AAACGTAGTTACTATCATCAAGTAGTTCTTTATCAAAAAGGAAAGAGATTACAAGGTTTCT 607
 Db 941 CTAATGGATTAAAGAAAGTAGTGGAAATATCAAGTAGTTCTTTATCCAGAAGAA 1000
 Qy 608 CAACCTTATAGAGCTACAAATTTAT---AATGAAGGAACCTGCATATGTTATATACACCACTAA 664
 Db 1001 AAAGAGTTCAACTAAGAGTGTCTGCAAGGCTTCAATTTGCTGTATGTAAGAAATTCACCAGTTA 1060

QY 665 TATTAACTCTAAATCTTACTAGTAAGAGTAATTTAAAGACTGCAGTAGAAGAGTTACAAA 724
DB |||||
QY 1061 AATTAACTCTTAAGTTCAGATAGAGAAAGACTTTAAAGAGTTATGTGGATGATTTAAGAA 1120
DB |||||
QY 725 AATTGAATCGTAGTTATTCTTAATCTAACCTTTAGCTGGTGAATGACAGAAATCAAAACAG 784
DB |||||
QY 1121 CATATTAATAAGATATTCAATGCTTATAGAAGTAGCAGGAGAAAGATAGAACTG 1180
DB |||||
QY 785 CTATAGAGATAGTAAGAGATATTACAATAATGATGGCGAGAAATCAGATCATTCAGCTG 844
DB |||||
QY 1181 CAATAGCATTAAGTCAAAATATTATAACTCTGATGAAAA-----TGCTA 1228
DB |||||
QY 845 ATGCTTAAAGAGATGTTAAAAATGTTGTATTAGTAGGTGCAATGCACTAGTAGATGAT 904
DB |||||
QY 1229 TATTAGAGATTCAGTTGATATGATGATTTGTTGGAGGAATGCAATAGTTGATGGAC 1288
DB |||||
QY 905 TAGTTGCGGCTCTTTAGCAGCAGAAAGAGATGCTCCACTATATTATTAATCTTCAAAAGATA 964
DB |||||
QY 1289 TTGTAGCTTCTCTTTAGCTTCTGAAAAGAAAGCTCTTTATTATTAACTTCAAAAGATA 1348
DB |||||
QY 965 AATTAGATTCGTAGTAAATCTGAATTAAGAGAGTTTGTAGACTTAAAACTTCAACAG 1024
DB |||||
QY 1349 AATTAGATTCAGCGTAAAGCTGAATTAAGAGAGTTATGAATATAAAGAGTACAAACAG 1408
DB |||||
QY 1025 AAGTA--ACAGGAAAAACAGTTTATATAGTCTGCTGAGTGAATAGTGTATCTAAAGAA 1081
DB |||||
QY 1409 GTATAAATCTTCAAGAGAAAGTTTATTAGCTGGTGGAGTTAATTTCTATCTTAAGAAG 1468
DB |||||
QY 1082 TTGTAAACAGAAATTAGAATCAATGGGATTAAGAGTTGAAGAGATTTCTCAGGTGATGATGAT 1141
DB |||||
QY 1469 TAGAAATGAATTAAGAGATATGGGACTTAAAGTTACAGATTAGCAGGAGATGATAGAT 1528
DB |||||
QY 1142 ATGAACCTCTTTAAAAATAGCAGGTGAATAGCTTAGAATATGATAGGCTTATGTAG 1201
DB |||||
QY 1529 ATGAACCTCTCTAAAAATAGCTGATGAAGTAGGTCTTGTATATGATTAAGCAATTTGTAG 1588
DB |||||
QY 1202 TTGCTGGAACAGATTAGCAGATGCCATGATGATAGCTTCCAGTTGCTTCTACTAAATTAG 1261
DB |||||
QY 1589 TTGAGGAGCAGGATTAGCAGATGCCATGATGATAGTATAGTCCAGTTGCATCTCAATTAGAA 1648
DB |||||
QY 1262 ATGCTAATGCTGTTGTAGATAGAACAAATGGACATGCTACTCCAAATAGTTGTTGTAGATG 1321
DB |||||
QY 1649 ATGCTAATGTTAAATGCAATTTAGCTGATGCTGATGCTACACCAATAGTAGTTGTAGATG 1708
DB |||||
QY 1322 GAAAGCTGATAAATATCTGATGACTTAGATAGTTTCTTAGGAGCGCTGATGATAGATA 1381
DB |||||
QY 1709 GAAAGCTAAACTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1768
DB |||||
QY 1382 TAATAGCTGATTTGCAAGTGTATCTCAAAAGATGGAAGAGCTATATCAGATGCTACTG 1441
DB |||||
QY 1769 TAATAGCTGAGAAAAACAGTGTATCTAAAGATGTTGAATATGCAATAGATGATGCTACAG 1828
DB |||||
QY 1442 GTAAGCGGTTACAAAGAGTTAAAGCGCAGATAGACAGACACTAACTCTGAAGTTATAA 1501
DB |||||
QY 1829 GTAAATCTCCAGATAGATATAGTGGAGTATAGACAGCAAGCACTAATGCAAAAGTTATAA 1888
DB |||||
QY 1502 AAACATATTATGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
DB |||||
QY 1889 AAGAATCTTCTTATTATCAAGATAAATTAATATATGATTAATAAAGAGTAGTTAATTTCTTTG 1948
DB |||||
QY 1562 GTGCTTCAAGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1621
DB |||||
QY 1949 TAGCTAAGAGATGGTTCTACTAAAGAGATCAATTAGTTGATGCTTTTAGCAGCAGCTCCAG 2008
DB |||||
QY 1622 A-----AGAGATCAATTAGTTG 1639
DB |||||
QY 2009 TTGAGGCAAACTTTGGTGAATCTTAATCTGATGTTAGCCAGTAGATAGATAAGATGCTA 2068
DB |||||
QY 1640 ATGCATTAGCAGTAGGAGCTGTTGCTGGATATAA-----CTTGCTCCAGTTGAT 1690
DB |||||
QY 2069 AAGTATTAACTGGTTCTGATAATGATAAATAAATTAGTATCTCCAGCACCTATAGTAT 2128
DB |||||
QY 1691 TAGCTACTGATCTTCTTCTGATCAATCGTTGCTGATATAGCAAGAGTTGCTAGAGAAA 1750
DB |||||

DB 2129 TAGCTACTGATCTTTATCTTTCAGATCAAGAGTGATCTATATAGTAAAGTTCTTGATAAG 2188
QY 1751 AATATTTCTAAAGATTTAAACAAAGTTGGTCAAGGAATAGCTAAATTCAGTTATAAACA 1810
DB |||||
QY 2189 AATATGGGAGAAACTTAGTTCAAGTTGGTAAAGGTATAGCTACTTCCAGTTATAAACA 2248
DB |||||
QY 1811 TGAAGAGATTTATTAGATATG 1830
DB |||||
QY 2249 TGAAGAGATTTATTAGATATG 2268
DB |||||
RESULT 4
ABS65378
ID ABS65378 standard; DNA; 2271 BP.
XX
AC ABS65378;
XX
XX 15-NOV-2002 (first entry)
XX Clostridium difficile strain 171448 slpA DNA sequence.
XX
XX Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
KW C. difficile infection; immune response; C. difficile associated disease;
KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
KW C. difficile vaccine; antibacterial; slpA; ds.
XX
OS Clostridium difficile.
XX
XX WO200262379-A2.
XX
XX 15-AUG-2002.
XX
XX 11-FEB-2002; 2002WO-IE000017.
XX
XX 09-FEB-2001; 2001IE-00000137.
XX
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
XX Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
PI
XX
XX WPI; 2002-643389/69.
XX
XX Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
PT associated diseases, comprises Clostridium difficile genes or
PT peptide/polypeptides or its derivative, fragment, mutant or variant.
XX
XX Claim 14; Page 73-74; 85pp; English.
XX
XX The present invention relates to a new vaccine for treatment/prophylaxis
CC of Clostridium difficile associated disease. The vaccine of the invention
CC comprises C. difficile gene or peptide/polypeptide or derivative,
CC fragment, mutant or variant of the peptide/polypeptide which is
CC immunogenic in humans or to which immunoreactivity is detected in
CC individuals who have recovered from C. difficile infection. The molecules
CC of the invention are used in prophylaxis or treatment of C. difficile
CC associated disease, where the medicament is a vaccine. The invention can
CC be used by administering the vaccine preparation to a host to raise an
CC immune response. The invention is further useful in the preparation of a
CC medicament for treatment or prophylaxis of C. difficile infection or C.
CC difficile associated disease. The invention is also useful in passive
CC immunotherapy for established C. difficile infection, for the eradication
CC of C. difficile associated disease, or for passive vaccination of an
CC individual with C. difficile infection. Interleukin is useful as an
CC adjuvant in C. difficile vaccine. The present nucleic acid sequence
CC represents a Clostridium difficile slpA DNA sequence of the invention
XX
SQ Sequence 2271 BP; 895 A; 263 C; 443 G; 670 T; 0 U; 0 Other;

Query Match 29.68; Score 542; DB 6; Length 2271;
Best Local Similarity 59.48; Pred. No. 4.4e-90;
Matches 1117; Conservative 0; Mismatches 670; Indels 93; Gaps 8;

Key	Location/Qualifiers	Query
CDS	1..2271	
FT	/*tag= a	
FT	/product= "S-layer protein"	
XX	WO200173040-A1.	
XX		
XX	04-OCT-2001.	
XX		
XX	23-MAR-2001; 2001WO-GB001305.	
XX		
XX	24-MAR-2000; 2000GB-00007263.	
XX	(UNLO) IMPERIAL COLLEGE SCI TECHNOLOGY & MED.	
XX		
XX	Fairweather NF, Calabi E;	
XX		
XX	WPI; 2001-616508/71.	
XX	P-PSDB; AAG79242.	
XX		
XX	Novel polypeptides and polynucleotides of cell wall proteins of	
XX	Clostridium difficile especially S-layer cell wall protein useful for	
XX	preventing and treating the infection caused by the bacteria.	
XX		
XX	Claim 3; Page 57-58; 62pp; English.	
XX		
XX	The present sequence encodes a surface layer (S-layer) protein of	
XX	Clostridium difficile. The S-layer proteins are the predominant cell wall	
XX	protein. There are two distinct S-layer proteins in C. difficile, a 45	
XX	kDa and 36 kDa protein. S-layer polypeptides and polynucleotides are	
XX	useful for treating and/or preventing a disease associated with	
XX	C. difficile infection in a subject. Such diseases include	
XX	pseudomembranous colitis (PMC) in humans characterized by diarrhoea, a	
XX	severe inflammation of the colonic mucosa, and formation of	
XX	pseudomembranes that are composed of fibrin, mucus, necrotic epithelial	
XX	cells and leukocytes; gastrointestinal illness, abscesses, wound	
XX	infections, osteomyelitis, urogenital tract infections, septicemia,	
XX	peritonitis, and pleuritis	
XX		
XX	Sequence 2271 BP; 895 A; 263 C; 442 G; 671 T; 0 U; 0 Other;	
XX		
XX	Query Match 29.3%; Score 535.6; DB 4; Length 2271;	
XX	Best Local Similarity 59.2%; Pred. No. 6.6e-89;	
XX	Matches 1113; Conservative 0; Mismatches 674; Indels 93; Gaps 8;	
QY	32 CAGCTGTTACTGTAGTGTCTGCTCCAGTTTTCGACGAGCTTCAGATCTAATAT 91	
DB	401 CAGCAGATGCTATAATGCTGGAAATCTTCAGCAGATGGTGTTGTTACAAATCTGGAG 460	
QY	92 CACTACAAGATGGTCAAAATGATAGTATACAGTATCAATATCAATATCAAGCTAGTGCATTAG 151	
DB	461 CTGCTAGTGGTCTACTAGACAAATTCAGCAGGACAAATCTTGCATGCTGACTATTT 520	
QY	152 TAAAGGATATTTAGCAGCACAAAATTTAAACACAGGTGCGAGTTATTTTGAACAAAGATA 211	
DB	521 TTGACACAGCATATACAGATTTCATCTGAAACTGCGGTTTAAAGTTTACATATAAAGCAGATA 580	
QY	212 CAAAAGTTACT---TTCTATGATGCAATGAGAAAGATCTTCAACTCCAACTGGAGATA 268	
DB	581 TGAATGATACTAAATTTGGTAAAGCAGGTGAGACCAACTTATTCACCTGGGCTTACATTTG 640	
QY	269 AAAAGTTTATTCAGBACAACTTTAACTACAGCTTAATGGAATGAAGATTTATGTAAGA 328	
DB	641 AGATGGGTCTACAGAAAAAATTTGTTAAATTAGGGGACAGTGAATTTATATAGATATACTA 700	
QY	329 CAACTTTAAAA---AATTATGATGAGGAGATATGCTATTATAGATTTAACTTTATAATA 385	
DB	701 AAGCTCTTAACTTACTGTGTTCTGGAGTAAAGCAACTGTTAAAGTTTGCCTGAAAAA 760	
QY	386 ATGCTAAAACTGTG-----AAATTTAAAGTAGTACGACTAGTGA 427	
DB	761 CACCAAGTGCCAGTGTTCACACAGTAATAACAAAGCTTGAATAATTAATGCTTAAAGNAG 820	

DB 1889 AGAATCTCTTATTATCAGATAACTTAATAATGATAAAGAGTAGTTATTTCTTTG 1948
QY 1562 GTGCTTCAAGTAGTGATGAGGAGTATTTAATTTCTATGAGTAAAGATGATCTACAA 1621
DB 1949 TAGCTAAAGATGGTCTTACTAAAGAGATCAATTAGTTGATGCTTTAGCAGCAGCTCCAG 2008
QY 1622 A-----
DB 2009 TTGAGCAAACTTTGGTGAACCTTTAATTTCTGATGAGGTAAGCCAGTAGATAAGATGGTA 2068
QY 1640 ATGCATTAGCAGTAGGAGCTGTGCTGGATATAA-----CTTGCTCCAGTTGTAT 1690
DB 2069 AGATTAATGCTGTTCTGATATGATAAATAAATAGTATCTCAGCAGCTATAGTAT 2128
QY 1691 TAGCTACTGATCTTTATCTTCTGATCAATTCGGTCTGATAGCAAAAGTTGTAGGAGAA 1750
DB 2129 TAGCTACTGATCTTTATCTTCTGATCAAAAGTGATCTATAAGTAAAGTTCTTGATAAG 2188
QY 1751 AATATTTCTAAGATTTTAAACAGTTGGTCAAGGATAGCTAATTCAGTTTATAACAAA 1810
DB 2189 ATAATGAGAGAACTTTAGTTCAAGTTGGTAAAGGTATAGCTACTTTCAGTTTATAACAAAT 2248
QY 1811 TGAAGATTTATTAGATATG 1830
DB 2249 TAAAGATTTATTAGTATG 2268

RESULT 6
ABS65372
ID ABS65372 standard; DNA; 2217 BP.
XX AC ABS65372;
XX DT 15-NOV-2002 (first entry)
XX DE Clostridium difficile strain 173644 slpA gene.
XX KW Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
KW C. difficile infection; immune response; C. difficile associated disease;
KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
KW C. difficile vaccine; antibacterial; slpA; gene; ds.
XX OS Clostridium difficile.
XX PN WO200262379-A2.
XX PD 15-AUG-2002.
XX PF 11-FEB-2002; 2002WO-IE000017.
XX PR 09-FEB-2001; 2001IE-00000137.
XX PA (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX PI Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
XX WPI; 2002-643389/69.
XX P-PSDB; ABG80351.
XX DR Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
XX associated diseases, comprises Clostridium difficile genes or
XX peptide/polypeptides or its derivative, fragment, mutant or variant.
XX PS Claim 49; Page 45-48; 85pp; English.
XX CC The present invention relates to a new vaccine for treatment/prophylaxis
CC of Clostridium difficile associated disease. The vaccine of the invention
CC comprises C. difficile gene or peptide/polypeptide or derivative,
CC fragment, mutant or variant of the peptide/polypeptide which is
CC immunogenic in humans or to which immunoreactivity is detected in
CC individuals who have recovered from C. difficile infection. The molecules
CC of the invention are used in prophylaxis or treatment of C. difficile

CC associated disease, where the medicament is a vaccine. The invention can
CC be used by administering the vaccine preparation to a host to raise an
CC immune response. The invention is further useful in the preparation of a
CC medicament for treatment or prophylaxis of C. difficile infection or C.
CC difficile associated disease. The invention is also useful in passive
CC immunotherapy for established C. difficile infection, for the eradication
CC of C. difficile associated disease, or for passive vaccination of an
CC individual with C. difficile infection. Interleukin is useful as an
CC adjuvant in C. difficile vaccine. The present nucleic acid sequence
CC encodes a Clostridium difficile slpA protein, as described in the
XX invention
SQ Sequence 2217 BP; 856 A; 271 C; 432 G; 658 T; 0 U; 0 Other;
Query Match 27.5%; Score 503.4; DB 6; Length 2217;
Best Local Similarity 59.6%; Pred. No. 5.3e-83;
Matches 950; Conservative 0; Mismatches 616; Indels 27; Gaps 5;
QY 250 TCAACTCCAACTGGAGATAAAGCTTTATTCCAGAACAACTTTAACTACAGCTAATGGA 309
DB 640 TCTGTTCTCTACTACAGGCTTAACATTAACCTGCTGATACAACTGCAACACAGATGTAAT 699
QY 310 AATGAAGATTATGTAAAGACAACTTTAAAAAATTTAGATGCAGGAGAAATATCTATTATA 369
DB 700 ATTTCTGATGTTATGAGTGCAITTTAAATTTAATGCTACTGATACGATTTGATTTCCCA 759
QY 370 GATTTAACTTATAATAATGCTTAAACTGTTGAAATTTAAAGTAGTAGAGCTAGTGAAAA 429
DB 760 GCTGGTTCATCAGCTTCTACTCTTAGAGCAAGTATAAAGTAAATAATGCAAAAGAGAA 819
QY 430 ACAGTAGTTGTTATCTAGTGATCGGAAATAATAGTGCAGGAGATATAGCTGAAAAATATG 489
DB 820 TCTATAGATGTTGATTCAAGTTTCAATAGAACAGCTGAAGATTTAGCTGAAAAATATGTA 879
QY 490 TTTGAAGACAAAGACTTAGAAAAATGCACCTAAAAAATATAAAATGCCTCAGATTTC----- 543
DB 880 TTTAAACCAAGAGATGTGAATAAAACTTATGAGGCACTGACTGATTTATATAAAGAGGT 939
QY 544 AGTAAAACTGATGTTATCTATCAAGTAGTTCTTTATCCAAAAAGGAAAGAGATTACAAGGT 603
DB 940 ATAAACAAGTAATCTTATCACTCAAGATGTTGGAATAATCAAGTTGTTTATTGCTCAA 999
QY 604 TTCTCACTTATAGAGCTACAAATTTAATGAAGCACTGTCATATGTTAATACACAGTA 663
DB 1000 GGAAGAGATTAACCTACTAAAGGACAACTTGAACCTTTAGCAGATGAATAATTTCTCTCT 1059
QY 664 ATATTAACTCTAAAAATCTACTAGTAGAGTAAATTTAAAGACTGCAGTAGAGAGTTACAA 723
DB 1060 AAGGTAAACAATAAAGCAGATAAAGTAAAGACTTTAAAGATTTATGTTGAAGATTAAAA 1119
QY 724 AAATTGAATGCTAGTTATTTCTAATACTACAACTTTAGCTGGTGATGACAGAAATACAAACA 783
DB 1120 AATGCTAAACAATGGATATTTCAAAATTTCTGTTGTTAGCAGGTTGAAGATAGATAAACA 1179
QY 784 GCTATAGAGATAAGTAAAGATTTTACAAATATGATGCGGAGAAATCAGATCATTCAGCT 843
DB 1180 GCAATAGAGTTAAGTAGCAAAATACATATACTCTGATGATGACAA-----TGCA 1227
QY 844 GATGTTAAAGAGAAATGTTAAAAATCTGTATTAGTAGGTGCAAAATGCACCTAGTAGTGA 903
DB 1228 ATAATTAAGATCCAGTTTAAACATGTTGTTTGTGTTGTTCTCAAGCTGTAGTTGATGG 1287
QY 904 TTAGTGGGGTCCCTTTTAGCAGCAGAAAGATGTCCTCACTATTTATTAATCTTCAAAAGAT 963
DB 1288 CTTGTAGCTTCACCTTTAGCATCTGAAAAAAGAGCTCTCTTTACTATTAACTTACAGCAGGA 1347
QY 964 AAATTAGATTCCTCAGTAAATCTCAAAATGAAGAGAGTTTATAGCTTAAAACTTCAACA 1023
DB 1348 AAATTAGATTCAGTTGTTAAAGCTGAGTTGAAAGAGATTAAGTATTTAAATCTCAACA 1407
QY 1024 GAAGTA---ACAGGAAAAACAGTTTATATAGCTGTTGGAGTTTAAATAGTGTATCTAAAGAA 1080
DB 1408 GGTGTAAATACTTCTTAAAAAAGTTTACTTTAGCTGTTGGAGTAAACTCTATATCTAAGAT 1467

PD	15-AUG-2002.
XX	11-FEB-2002; 2002WO-IE000017.
PF	09-FEB-2001; 2001IE-00000137.
XX	(QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX	Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
PI	WPI; 2002-643389/69.
XX	Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
PT	associated diseases, comprises Clostridium difficile genes or
PT	peptide/polypeptides or its derivative, fragment, mutant or variant.
XX	Claim 16; Page 75-76; 85pp; English.
PS	The present invention relates to a new vaccine for treatment/prophylaxis
XX	of Clostridium difficile associated disease. The vaccine of the invention
CC	comprises C. difficile gene or peptide/polypeptide or derivative,
CC	fragment, mutant or variant of the peptide/polypeptide which is
CC	immunogenic in humans or to which immunoreactivity is detected in
CC	individuals who have recovered from C. difficile infection. The molecules
CC	of the invention are used in prophylaxis or treatment of C. difficile
CC	associated disease, where the medicament is a vaccine. The invention can
CC	be used by administering the vaccine preparation to a host to raise an
CC	immune response. The invention is further useful in the preparation of a
CC	medicament for treatment or prophylaxis of C. difficile infection or C.
CC	difficile associated disease. The invention is also useful in passive
CC	immunotherapy for established C. difficile infection, for the eradication
CC	of C. difficile associated disease, or for passive vaccination of an
CC	individual with C. difficile infection. Interleukin is useful as an
CC	adjuvant in C. difficile vaccine. The present nucleic acid sequence
CC	represents a Clostridium difficile slpA DNA sequence of the invention
XX	Sequence 2217 BP; 856 A; 271 C; 432 G; 658 T; 0 U; 0 Other;
SQ	
QY	Query Match 27.5%; Score 503.4; DB 6; Length 2217;
Db	Beat Local Similarity 59.6%; Pred. No. 5.3e-83;
QY	Matches 950; Conservative 0; Mismatches 616; Indels 27; Gaps 5;
Db	250 TCAACTCCAACTGGAGATAAAAAAGTTTATTTCAGAACAACTTTAACTACAGCTAATGGA 309
QY	310 AATGAAGATTTATGTAAGACAACTTTTAAAAATTTTAGATGTCAGGAGAAATATGCTATTATA 369
Db	700 ATTTCTGATGTTATGATGCGATTAAATTTAATGGTACTGATACGATTTAGTGGATTTCCCA 759
QY	370 GATTTAACTTATAATAATGCTTAAACCTGTTGAAATTTAAAGTAGTAGCAGCTAGTGAAAA 429
Db	760 GCTGGTTCATCAGCTTCTACTCTTTAGAGCAAGTATAAAGTAAATGCAAAAGAAAGAA 819
QY	430 ACAGTAGTTGTATCTAGTATGTCGAAAAATAGTGGCAAAAGATATAGCTGAAAAATATGTG 489
Db	820 TCTATAGATGTTGATTCAGTTTCAATGAAACAGCTGCAAGATTTAGCTGAAAAATATGTA 879
QY	490 TTTGAAGCAAAAGACTTTAGAAAAATGCATAAAAACTATAAATGCTCCTCAGATTTCTC 543
Db	880 TTTAAACCAAGAGATGTGATATAAACTTATGAGGCACTGACTGATTTATATTAAGAAGGT 939
QY	544 AGTAAAACTGATAGTTACTATCAAGTAGTTCTTTTATCCAAAAGGAAAGAGATTACAAAGT 603
Db	940 ATAAAGATTAATCTTATCACTCAAGATGGTGGAAAAATATCAAGTTGTTTATTGCTCAA 999
QY	604 TTCTCAACTTATAGAGCTACAAATTTATAATGAAGGAACCTGATATGTAATATACACAGTA 663
Db	1000 GGAAGAGATTAATCTACTTAAAGGAGCAACTGGAACCTTTAGCAGATGAAAAATTTCTCTCT 1059
QY	664 ATATTAACTCTAAAAATCTACTAGTAAGAGTAATTTTAAAGACTGCGTAGAGAGTTTACAA 723
Db	1060 AAAAGTAAATAAAGCAGATTAAGTAAAAAGACTTAAAGAGATTATGTTGAAGATTAAAAA 1119

Clostridium difficile strain 173644 slpA DNA sequence.

Vaccine; Clostridium difficile; immunogenic; immunoreactivity; C. difficile infection; immune response; C. difficile associated disease; passive immunotherapy; passive vaccination; interleukin; adjuvant; C. difficile vaccine; antibacterial; slpA; ds.

Clostridium difficile.

WO200262379-A2.

```
QY 724 AAATTGAATGCTAGTATTCTTAATACTACAACTTTAGCTGGTGATGACAGAAATACAAACA 783
DB 1120 AATGCTAAACAATGGGATATTCAAAATCTGTTGTGAGCAGGTTGAAGATAGATAGAAACA 1179
QY 784 GCTATAGAGATAAGTAAGAAATATTACAAATATGATGGCGAGAAATCAGATCATTCAGCT 843
DB 1180 GCAATAGAGTTAAGTAGCAAAATCTATAAATCTGATGATGACA-----TGCA 1227
QY 844 GATGTTAAAGAGAATGTTAAAAATGTTGATAGTGGTGCAAAATGCACCTAGTAGAGGA 903
DB 1228 ATAACCTAAAGATCCAGTTAAACAATGTTGTTTGTAGTTGGTTCTCAAGCTGATGATGG 1287
QY 904 TTAGTTGGGGCTCTTTAGCAGCAGAAAGAAAGATGCTCCACTATTATTAACTTCAAAAGAT 963
DB 1288 CTTGTAGCTTCACTTTAGTACTCTGAAGAAAGAGCTCTTTACTATTAACTTTAGCAGGA 1347
QY 964 AAATTAGATTGCTCAGTAAAAATCTGAAATAAAGAGAGTTTATAGACTTAAAGAACTTCAACA 1023
DB 1348 AAATTAGATTCAAGTGTAAAGCTGAGTTGAAAGAGTAATGGATTTAAATCTCAACA 1407
QY 1024 GAAGTA---ACAGGAAAAACAGTTTATATAGCTGGTGGAGTTAATAGTGTATCTAAAGAA 1080
DB 1408 GGTGTAATACTCTTAAGAAAGTTTACTTTAGCTGGTGGAGTAAACTCTATACTTAAGAT 1467
QY 1081 GTTGTAAACAAGATTAGAAATCAATGGGATTTAAAGTTGAAAGATTTCTCAGGTGATGATGA 1140
DB 1468 GTAGAAATGAATTAAGAGATATGGGACTTAAAGTTTACAAGATTTATCAGGAGATGATGA 1527
QY 1141 TATGAACTTCTTTTAAAAATAGCAGTAAATAGGCTTAGATAAATGATAGGCTTATGTA 1200
DB 1528 TATGAACTTCTTTTAACTTATAGCTATAGCTGATGAAATAGGCTTTGATTAATGATTAAGCTTTGTA 1587
QY 1201 GTTGTGGAACAGGATTAGCAGATGCCATGATATAGCTTTCAGTTGCTTCTCTAAATTA 1260
DB 1588 GTTGAGGAACAGGATTAAGCGGATGCTATGATGATAGCTCCAGTTGCTTCTCAATTAAGA 1647
QY 1261 GATGTAATGTTGTTGATAGATAGAACAAATGGACATGCTACTCCAAATAGTTGTTGATAT 1320
DB 1648 AACTCAATGGAGAACTTGA---CTTAAAGGTGATGCAACTCCAATAGTAGTTGTTGAT 1704
QY 1321 GGAAGAGCTGATAAATACTGATGACTTAGATAGTTTCTTAGGAAGCGCTGATGATAG 1380
DB 1705 GGAAGAGCTGATAAATAAATCTGAAGTAAAGATTTCTTAGATGATTCACAAGTTGAT 1764
QY 1381 ATAATAGTGGATTTGCAAGTGTATCTGAAAAGATGGAAGAGCTATATCAGATGCTACT 1440
DB 1765 ATAATAGTGGTGTAAATAGTGTCTTAAGAGTAATGGAAGCAATAGATGATGCTACT 1824
QY 1441 GGTAAAGCGGTTACAAGAGTTTAAAGCGACGATAGACAAAGACACTAACTCTGAAGTTATA 1500
DB 1825 GGAAGATCACTCAGAGATATATAGTGGAAGATAGACAAAGCAAAATGCTTAAAGTTATA 1884
QY 1501 AAAACATATTATGCTAATGATGCTGAAATAGCTTAAAGCTGCAAGTTTATAGATAAGATTCA 1560
DB 1885 AAAGAGATGATTTCTTTTAAATGGAAGATTTACAACTTCTTTGATGATTAAGATGT 1944
QY 1561 GGTGCTTCAAGTAGTGATGACAGAGTATTAAATTTCTATGCTAAGATGATGATCTACA 1620
DB 1945 TCAACTTAAAGAGATCAATTAGTAGATGCTTTTAGCAGGTTGCTGCAATTCGTGTAACTTT 2004
QY 1621 AAAGAGATCAATTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
DB 2005 GGTGTAACAGTATAGATAATGAAGGAAACCTACAGTTGCTGATGATGATGATGATGATG 2064
QY 1681 CAGTTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 2065 CCAATGTTTAAAGCAACAGATTTCTTTATCTTCTGATCAATCGTTGCTATAGCAAGATT 2124
QY 1741 GTA---GGAGAAAAATTTCTAAAGATTTTAAACAAGTTGGTGAAGGAATAGCTAATTTCA 1797
DB 2125 GTAAATGATGACGCTAATACTAAGATCTAGTTTCAAGTTGGTAAAGGTATAGCTACTTCA 2184
```

```
QY 1798 GTTATACAAAAATCAAGATTTTATTATGATATG 1830
DB 2185 GTTGTAGTAAATAAAGATTTTATTAGATATG 2217

RESULT 8
AAI65839
ID AAI65839 standard; DNA; 2145 BP.
XX
AC AAI65839;
XX
DT 03-JAN-2002 (first entry)
XX
DE Nucleotide sequence of S-layer protein of C. difficile strain 17.
XX
KW Surface layer protein; S-layer protein; pseudomembranous colitis; PMC;
KW cell wall protein; gastrointestinal illness; abscess; wound infection;
KW osteomyelitis; urogenital tract infection; septicemia; peritonitis;
KW pleuritis; ss.
XX
OS Clostridium difficile.
XX
FH Key Location/Qualifiers
CDS 1..2145
FT /*tag= a
FT /product= "S-layer protein"
XX
PD WO200173040-A1.
XX
PF 04-OCT-2001.
XX
PR 23-MAR-2001; 2001WO-GB001305.
XX
PA 24-MAR-2000; 2000GB-00007263.
XX
PI (UNLO ) IMPERIAL COLLEGE SCI TECHNOLOGY & MED.
XX
PI Fairweather NF, Calabi E;
XX
DR WPI; 2001-616508/71.
DR P-PSDB; AAG79240.
XX
PT Novel polypeptides and polynucleotides of cell wall proteins of
PT Clostridium difficile especially S-layer cell wall protein useful for
PT preventing and treating the infection caused by the bacteria.
XX
PS Claim 3; Page 56-57; 62pp; English.
XX
CC The present sequence encodes a surface layer (S-layer) protein of
CC Clostridium difficile. The S-layer proteins are the predominant cell wall
CC protein. There are two distinct S-layer proteins in C. difficile, a 45
CC kDa and 36 kDa protein. S-layer polypeptides and polynucleotides are
CC useful for treating and/or preventing a disease associated with
CC C.difficile infection in a subject. Such diseases include
CC pseudomembranous colitis (PMC) in humans characterized by diarrhoea, a
CC severe inflammation of the colonic mucosa, and formation of
CC pseudomembranes that are composed of fibrin, mucus, necrotic epithelial
CC cells and leukocytes; gastrointestinal illness, abscesses, wound
CC infections, osteomyelitis, urogenital tract infections, septicemia,
CC peritonitis, and pleuritis
XX
SQ Sequence 2145 BP; 846 A; 261 C; 422 G; 616 T; 0 U; 0 Other;

Query Match 26.4%; Score 483.2; DB 4; Length 2145;
Best Local Similarity 62.7%; Pred. No. 2.7e-79;
Matches 911; Conservative 0; Mismatches 433; Indels 108; Gaps 6;

QY 394 ACTGTTGAATTAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTTGTATCTAGTATCG 453
DB 784 ACTATAAAGTTAGAGTTACAGTGCAGAAAGAAAGTAATCTATTGATGGATTCAAGTTCA 843
QY 454 AAAAAATAGTGCAGAAAGATATAGCTCAAAAAATATGTGTTTGAAGACAAAGACTTAGAAAT 513
```



```

844  TATATTAGTCTGCTCAAAATTTAGCTAAAGAAATATGTATTATTAATCCTAAAGAGGTTTCTGAA 903
QY
514  GCACATAAACTATATAATGCCTCAGATTTTCAGTAAAACTGATGTTACTATCAAGTAGTT 573
Db
904  GCTTATAATGCAATAGTTGCAATTAACAAATGATGGAATAGATCTGATTTAGTACAAATTA 963
QY
574  CTTTATCCAAAAGGAAAGAGATTAACAAGGTTTCTCAACTTTATAG--AGCTACAAATTA 630
Db
964  GTTAATGGAAATATCAAGTTATTTCCTATCCAGAGGAAAGATTTAGAAACTAAATCT 1023
QY
631  AATGAAGAACTGCATATGATATGATACACCGATTAATTTAACTCTAAAACTACTAGTAAG 690
Db
1024  GCAGATATATATGCTGATGAGATAGTCCAGCTAAATTAACATATAAAAGCTATAAATTA 1083
QY
691  AGTAATTTAAAGACTGCAGTAGAAGAGTTACAAAATTTGAATGCTAGTTATTCTAATACT 750
Db
1084  AAGAGTTTAAAGATTTATGTAGATGATTTTAAACACATACATTAATCTTACTCAATGTT 1143
QY
751  ACAACTTTAGCTGGTGATGACAGAAATCAAAACAGCTATAGAGATTAAGTAAGTAATTTAC 810
Db
1144  GTAACAGTAGCAGAGAGATAGATAGAACTGCTATAGAAATTAAGTAGTAATATTAT 1203
QY
811  AATAATGATGCGAGAAATCAGATCAATTCAGCTGATGTTAAAGAGAAATGTTAAAAATGTT 870
Db
1204  AATTCTGATGATAAAATGCAATAACTGATGATGCAAGTTAAT-----AATATA 1251
QY
871  GTATTAGTAGTGCAATGCACTAGTAGATGATGTTGCGGCTCCTTTAGCAGCAGAA 930
Db
1252  GTATTAGTTGGATCTACATCTATAGTTGATGGTCTTGTGTCATCACCATTAGTCTCAGAA 1311
QY
931  AAAAGATGCTCCACTATTATTAACTTCAAAAGATAAAATAGATTTCGTCAGTAAATCTGAA 990
Db
1312  AAAACAGCTCATTATTATTAACTTCAAAAGATAAATAGATTCAACAGTAAATCTGAG 1371
QY
991  ATAAGAGAGTTTATAGACTTAAAACTTCAACAGAGTA---ACAGAAAAAACAGTTTAT 1047
Db
1372  ATAAAAGAGTTATGAATTTAAAGAGTGATCTGTTATAAATATCTTCTAAAAAAGTTTAT 1431
QY
1048  ATAGCTGGTGGATTAATAGTGTATCTTAAAGAGTTGTTAAAGAAATTTAGAAATCAATGGGA 1107
Db
1432  TTAGCTGGTGGAGTTAAATCTTATATCTTAAAGATGTAGAAATGAATGAAAAATATGGGC 1491
QY
1108  TTTAAAGTTCGAAAGATTCAGGTGATGATAGATATGAAATCTCTTTTAAAAAATAGCAGGT 1167
Db
1492  CTTAAAGTTACTAGATTATCAGGAGAGACAGATACGAACTCTTTAGCAATAGCTGAT 1551
QY
1168  GAAATAGGCTTAGATAATGATAGGCTTATGTTAGTTGGTGGAAACAGGATTAGCAGATGCC 1227
Db
1552  GAAATAGGCTCTTGATAATGATAAGCATTTGTAGTTGGTGGTACTGCGATTAGCAGATGCT 1611
QY
1228  ATGAGTATAGCTTCAGTTGCTTCTTCTAATAATTAGATGTTAATGTTGTGTAGATAGACA 1287
Db
1612  ATGAGTATAGCTCCAGTTGCTTCT-----CAACTTAAA 1644
QY
1288  AATGGACATGCTACTCCAAATAGTTGTTGTAGATGAAAAAGCTGATAAAATATCTGATGAC 1347
Db
1645  GATGGAGATGCTACTCCAAATAGTTGTTGTAGATGAAAAAGCAAAAGAAATAGTATGAT 1704
QY
1348  TTAGATAGTTTCTTAGGAAGCGCTGATGTAGATATAATAGTGGATTGTTGCAAGTGTATCT 1407
Db
1705  GCTAAGAGTTTCTTAGGAACCTCTGATGTTGTATATAATAGTGGAAAAAATAGCGTATCT 1764
QY
1408  GAAAGATGGGAAGACTATATCAGATGCTACTGTTAAGCGTTACAGAGTTAAAGGC 1467
Db
1765  AAAGAGATTGAAGAGTCAATAGATAGTGCACCTGGAAGAACTCCAGATAGAAATAGTGGGA 1824
QY
1468  GACCATAGACAGACACTTAATCTGAAAGTTTATAAAACATATTATGCTAATGATCTGAA 1527
Db
1825  GATGACAGACAGCAACTAATGCTGAAAGTTTAAAA----- 1860
QY
1528  ATAGCTAAAGCTGACAGTTTATAGATAAAGATTCAGGTGCTTCAAGTAGTGATGCGAGGATA 1587
Db
1861 -----GAAGATGATTATTTCAAAAGATGGTGAAGTT 1890
```

```

1588  TTTAATTTCTATGTAGCTAAAGATGATCTACAAAGAAAGATCAATTAGTTGATGCAATTA 1647
Db
1891  GTGAATTTACTTTGTTGCAAAAGATGGTTCTACTAAAGAAAGATCAATTTAGTAGTCAATTA 1950
QY
1648  GCAGTAGGAGCTGTTGCTGG-----ATATAAACTTGTCTCCAGTTCTTATTAGCTACT 1698
Db
1951  GCAGCAGCACCAATAGCAGGTAGATTTAAGGAGTCTCCAGCTCCCAATCATACTAGCTACT 2010
QY
1699  GATTCCTTTATCTTCTGATCAATCGGTTGCTATAAGCAAAAGTTGTAGGAGAAAAATATTCT 1758
Db
2011  GATACTTTATCTTCTGACCAAAATGTAGCTGTAGTAAAGCAGTTCTCTAAAGATGCTGGA 2070
QY
1759  AAAGATTTTAAACAAGTTGGTCAAGAATAGCTAATTCAGTTTATTAACAAATGAAGAT 1818
Db
2071  ACTAACTTAGTTTCAAGTAGGTAAAGTATAGCTTCTTCAAGTTATTAACAAATGAAGAT 2130
QY
1819  TTATTAGATATG 1830
Db
2131  TTATTAGATATG 2142

RESULT 9
ABS65371
ID ABS65371 standard; DNA; 2145 BP.
XX
AC ABS65371;
XX
DT 15-NOV-2002 (first entry)
XX
DE Clostridium difficile strain 171862 slpA gene.
XX
KW Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
KW C. difficile infection; immune response; C. difficile associated disease;
KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
KW C. difficile vaccine; antibacterial; slpA; gene; ds.
XX
OS Clostridium difficile.
XX
WO200262379-A2.
PD
XX 15-AUG-2002.
XX
XX 11-FEB-2002; 2002WO-IE000017.
XX
XX 09-FEB-2001; 2001IE-00000137.
XX
XX (QUEB-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
XX Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
XX
XX WPI; 2002-643389/69.
XX
XX P-PSDB; ABG80350.
XX
XX Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
XX associated diseases, comprises Clostridium difficile genes or
XX peptide/polypeptides or its derivative, fragment, mutant or variant.
XX
XX Claim 48, Page 41-44; 85pp; English.
XX
XX The present invention relates to a new vaccine for treatment/prophylaxis
XX of Clostridium difficile associated disease. The vaccine of the invention
XX comprises C. difficile gene or peptide/polypeptide or derivative,
XX fragment, mutant or variant of the peptide/polypeptide which is
XX immunogenic in humans or to which immunoreactivity is detected in
XX individuals who have recovered from C. difficile infection. The molecules
XX of the invention are used in prophylaxis or treatment of C. difficile
XX associated disease, where the medicament is a vaccine. The invention can
XX be used by administering the vaccine preparation to a host to raise an
XX immune response. The invention is further useful in the preparation of a
XX medicament for treatment or prophylaxis of C. difficile infection or C.
XX difficile associated disease. The invention is also useful in passive
XX immunotherapy for established C. difficile infection, for the eradication
```


peptide/polypeptides or its derivative, fragment, mutant or variant.
 Claim 17; Page 77-78; 85pp; English.
 The present invention relates to a new vaccine for treatment/prophylaxis of Clostridium difficile associated disease. The vaccine of the invention comprises C. difficile gene or peptide/polypeptide or derivative, fragment, mutant or variant of the peptide/polypeptide which is immunogenic in humans or to which immunoreactivity is detected in individuals who have recovered from C. difficile infection. The molecules of the invention are used in prophylaxis or treatment of C. difficile associated disease, where the medicament is a vaccine. The invention can be used by administering the vaccine preparation to a host to raise an immune response. The invention is further useful in the preparation of a medicament for treatment or prophylaxis of C. difficile infection or C. difficile associated disease. The invention is also useful in passive immunotherapy for established C. difficile infection, for the eradication of C. difficile associated disease, or for passive vaccination of an individual with C. difficile infection. Interleukin is useful as an adjuvant in C. difficile vaccine. The present nucleic acid sequence represents a Clostridium difficile slpA DNA sequence of the invention

Sequence 2145 BP; 844 A; 261 C; 424 G; 616 T; 0 U; 0 Other;
 Query Match 26.2%; Score 480; DB 6; Length 2145;
 Best Local Similarity 62.6%; Pred. No. 1e-78;
 Matches 909; Conservative 0; Mismatches 435; Indels 108; Gaps 6;

394 ACTGTTGAAATTAAGTAGTAGCAGCTAGTGAAGAAACAGTAGTTCTATCTAGTGGCG 453
 784 ACTATAAAGTTAGAGTTACAGTGCAAGTGAAGAAAGAGTAATCTATTGATGGATTCAAGTTCA 843
 454 AAAAATAGTGCAAAAGATATAGCTGAAATATATGTGTTGCAACACAAAGACTTAGAAAT 513
 844 TATATTAGTCTGAAATTTAGCTAAATAATATGTATTTAATCCTAAAGAGGTTTCTGAA 903
 514 GCACATAAACTATAATGCTCAGATTTTCAGTAAACCTAGTACTATCAATCAAGTAGTT 573
 904 GCTTATAATGCAATAGTTGCAATTAACAATATGGAATAGATCTGATTTAGTACAATTA 963
 574 CTTTATCCAAAGGAAGAGATTACAAGGTTTCTCAACTTATAG---AGTACAATATAT 630
 964 GTTAATGGAATATCAAGTTATTTCTATCCAGAGGAAAGAGATTAGAACTAAATCT 1023
 631 AATGAAGAACTGCATATGTAATACACAGTAAATATTAATCTTAAATCTACTAGTAAG 690
 1024 GCAGATATAATAGCTGATGAGAGATAGTCCAGCTAAATAACTATAAAGCTAAATAATTA 1083
 691 AGTAATTTAAAGACTGCAGTAGAAGAGTTACAAAATTGAAATGCTAGTTATTTCTAATACT 750
 1084 AAAGATTAAAGATTATGTAGATGATTTAAACATACATATTAATCTACTCAATGTT 1143
 751 ACACTTTAGCTGTGTATGACAGAAATCAAAACAGCTATAGAGATAAGTAAGAAATATTAC 810
 1144 GTAACAGTACAGGAGAGATAGATAGAACTGCTATAGAAATTAAGTAAATATATAT 1203
 811 AATAATGATGGGAGAAATCAGATCATTCAGCTGATGTTAAAGAGATGTTAAATATGTT 870
 1204 AATCTCATGATAAAATGCAATTAACCTAGTATGATGCTGAGTCTTCTTAAATATA 1251
 871 GTATTAGTGGTCAATATGCACTAGTAGATGAGTATGTCGGCTCCTTTAGCAGAGAA 930
 1252 GTATTAGTGGATCTACATCTATAGTTGATGGTCTTGTGCAATCACCATTAGCTTCAGAA 1311
 931 AAAGATGCTCCACTATTTAATTAACCTTCAAAAGATAAATTAGATTGCTCAGTAAATCTGAA 990
 1312 AAAACAGCTCCATTTATTTAATCTCAAAAGATAAATTAGATTTCATCAGTAAATCTGAG 1371
 991 ATAAGAGAGTTTATGACTTAAACCTTCAACAGAGTA---ACAGGAAACACGTTTAT 1047
 1372 ATAAAAGAGTTATGAATTAAGAGTGATAGTCTGGTATAAATACTTCTTAAAGAGTTTAT 1431
 1048 ATAGCTGGTGGATTATAGTGTATCTAAAGAGTTGTTAACAGAGTTAGAAATCAATCGGA 1107

Db 1432 TTAGCTGGTGGAGTTAAATTTCTATATCTAAAGATGTAGAAGATGAATTTGAAAATATGGGC 1491
 Qy 1108 TTAAAAGTTTGAAGATTTCTCAGGTGATGATAGATATGAACTTTCTTTTAAAAATAGCAGGT 1167
 Db 1492 CTTAAAGTTTACTAGATTATCAGAGAGACAGATACGAACTTTCTTTAGCAATAGCTGAT 1551
 Qy 1168 GAAATAGGCTTTAGATAAATAGTAAGGCTTTATGTAGTTGGTGGAAACAGGATTAGCAGATGCC 1227
 Db 1552 GAAATAGGCTTTGATAAATAGTAAGGCAATTTGTAGTTGGTGGTACTGGATTGGCAGATGCT 1611
 Qy 1228 ATGATATAGCTTCAAGTTGCTTTCTACTATAAATTAGATGGTAAATGGTGTCTAGATAGCA 1287
 Db 1612 ATGATATAGCTTCAAGTTGCTTTCTACTATAAATTAGATGGTAAATGGTGTCTAGATAGCA 1644
 Qy 1288 AATGACATGCTTACTCAATAGTTGTTGTAGATGGAAGAGCTGATAAAATATCTGATGAC 1347
 Db 1645 GATGAGATGCTTACTCAATAGTTGTTGTAGATGGAAGAGCAAAAGAAATAGTATGAT 1704
 Qy 1348 TTAGATAGTTTCTTAGGAAAGCCTGATGATAGATATAATAGTGGATTTGCAAGTGTATCT 1407
 Db 1705 GCTAAGAGTTTCTTAGGAACTTCTGATGTTGATATAATAGTGGAAAAAATAGCGTATCT 1764
 Qy 1408 GAAAGATGGAAGAGCTATATCAGATGCTACTGTTAAAGGCGTTACAGAGTTAAAGGC 1467
 Db 1765 AAAGAGATTGAAGAGTCAATAGATAGTGCACACTGGAAAAAACCAGATAGAATAAGTGA 1824
 Qy 1468 GAGATAGACAGACACACTAACTCTGAAGTTTATAAAAAACATATATGCTAATGATGACTGAA 1527
 Db 1825 GATGACAGACAGCAACTAACTGCTGAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 1860
 Qy 1528 ATAGCTAAAGCTGCAGTTTATAGATAAGAGATTCAAGTGTCTCAAGTAGTATGATGAGAGTA 1587
 Db 1861 -----GAAGATGATTATTTCAAGAGATGCTGAAGTT 1890
 Qy 1588 TTTAATTTCTATGCTAAAGATGATCTACAAAAGAGATCAATTAGTTGATGCAATTA 1647
 Db 1891 GTGAATTTCTTTGTCAAAAGATGTTCTTCTAAAGAGATCAATTAGTATGATGCAATTA 1950
 Qy 1648 GCAGTAGGAGCTGCTGCTG---ATATAAACTTGTCTCAGTTGCTATTAGCTACT 1698
 Db 1951 GCAGCAGCACCAGATGAGTATTAAGGAGTCTCAGCTCAATCATCTAGCTACT 2010
 Qy 1699 GATTTCTTTATCTTCTGATCAATCGTTGCTATAGCAAAAGTTGTAGAGAGAAAAATATCT 1758
 Db 2011 GATCTTTATCTTCTGACCAAAATGTAGCTGTAAAGTAAAGCAGTTCTCTAAAGATGGTGA 2070
 Qy 1759 AAGATTAAACAAAGTTGGTCAAGGAATAGCTAATTCAGTTTATAACAAATCAAGAT 1818
 Db 2071 ACTAATCTTAGTTCAAGTAGTAAAGGTATAGCTTCTTCTAGTTATAACAAATCAAGAT 2130
 Qy 1819 TTATTAGATATG 1830
 Db 2131 TTATTAGATATG 2142

RESULT 11
 ABS65369
 ID ABS65369 standard; DNA; 2157 BP.
 XX
 AC ABS65369;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Clostridium difficile strain 170324 slpA gene.
 XX
 KW Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
 KW C. difficile infection; immune response; C. difficile associated disease;
 KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
 KW C. difficile vaccine; antibacterial; slpA; gene; ds.
 XX
 OS Clostridium difficile.

QY 836 ATTTCAGCTGATGTTAAGAGAAATGTTAAATAATGCTGTTATAGTAGTGCAAAATGCACTAG 895
 Db 1232 AAATGCAATACTGATAAGCGGTTAATGATATAGTATTAGTTGGATCTACATCTATAG 1291
 QY 896 TAGATGGATTAGTTGGGCTCCCTTTAGCAGCAGAAAAAGATGCTCCACTATTATTAACTT 955
 Db 1292 TTGATGGTCTGTTGCATCACCATTAGCTTCAGAAAAACAGCTCCATTATTATTAACCTT 1351
 QY 956 CAAAGATAAATAGATGTCGTCAGTAAATCTGAAATAAAGAGAGTTTGTAGACTTAAAAA 1015
 Db 1352 CAAAGATAAATAGATTCATCAGTAAATCTGAAATAAAGAGAGTTATGAACTTAAAGA 1411
 QY 1016 CTTCAACAGAGTA---ACAGGAAACACAGTTTATATAGCTGTCGGAGTTAATAGTAT 1072
 Db 1412 GTGACACTGGTATAAATCTCTTAAAAAAGTTTATTAGCTGGTGAGTTAATCTTATAT 1471
 QY 1073 CTAAGAAGTTGTAAACAGAAATAGAAATCAATGGGATTTAAAGTTTGAAGATTCTCAGGTG 1132
 Db 1472 CTAAGAATGTAGAAATGAATGAAACACATGGGCTTTAAAGTTACTAGATTATCAGGAG 1531
 QY 1133 ATGATAGATATGAACTCTTTAAATAATAGCAGGTGAAATAGGCTTTAGATAATGATAAGG 1192
 Db 1532 AAGACAGATACGAACTCTTTAGCAATAGCTGATGAAATAGCTCTTGATAATGATAAAG 1591
 QY 1193 CTTATGCTGTTGGTGAACAGGATTAGCAGATGCCATGAGTATAGCTTCAGTTGCTTCTA 1252
 Db 1592 CATTTGATGTTGGTGTCTGGAATAGCAGATGCTATGAGTATAGCTCCAGTTGCTTCT- 1650
 QY 1253 CTAATATAGATGTAATGTTGTAGATAGAACAAATCGACATGCTACTCCAATAGTTG 1312
 Db 1651 -----CAACTTAAAGATGGAGTGTCTCTCCAATAGTAG 1684
 QY 1313 TTGTAGATGGAAGCTGATATAAATATCTGATGACTTATAGTATGTTCTTAGGAAGCGCTG 1372
 Db 1685 TTGTAGATGGAAGCAAAAGAAATAAGTGATGATCTTAAGAGTTTCTTAGGAACTTCTG 1744
 QY 1373 ATGTAGATATAGTGTGATTTGCAAGTGATCTGAAAGATGGAAGAGCTATATACAG 1432
 Db 1745 ATGTTGATATAATAGTGTGAAATAATAGCGTATCTTAAAGAGATTGAAGAGTCAATAGATA 1804
 QY 1433 ATGCTACTGGTAAAGCGTTTAAAGAGTTTAAAGCGCAGCATAGACAAAGACATTAACCTGT 1492
 Db 1805 GTGCAACTGMAAACTCCAGATAGATAAGTGGAGATGATAGACAAAGCACTAATGCTG 1864
 QY 1493 AGTTTATAAACAATATATGCTAATGATGATGACTGAAATAGCTAAAGCTGCAGTTTATAGATA 1552
 Db 1865 AAGTTTAAAGAAAGATGATTATTTTC----- 1890
 QY 1553 AAGNTTCAGGTGCTTCAAGTATGATGATGAGGAGTATTTAATTTCTATGTAGCTAAAGATG 1612
 Db 1891 -----ACAGATGGTGAAGTTGTGAATTTACTTTTGTGCAAAAGATG 1930
 QY 1613 GATCTACAAAAGAGATCAATAGTTAGTGCATTTAGCAGTAGGAGCTGTTGCTGG----- 1667
 Db 1931 GTTCTACTAAGAGATCAATAGTAGTAGTGCCTTAGCAGCAGCAACCAATAGCAGGTAGAT 1990
 QY 1668 ----ATATAAATCTGCTCCAGTTGATATAGCTACTGATTTCTTTATCTTCTGATCAATCGG 1723
 Db 1991 TTAAGGAGTCTCCAGCTCCAATCATACTAGCTACTGATATCTTATCTTCTGACCAAAATG 2050
 QY 1724 TTGCTATAGCAAGTTGTAGGAGAAATATCTTAAAGATTTAACACAGTTTGTGTCAG 1783
 Db 2051 TAGCTGTAAAGTAAAGCAGTTCTTAAAGATGGTGGAAGTAACTTATAGTTCAAGTAGGTAAG 2110
 QY 1784 GAATAGCTAATTCAGTTTATAAACAATAATGAAAGATTTATTATAGATATG 1830
 Db 2111 GTATACCTCTTCAGTTTATAAACAATAATGAAAGATTTATTATAGTATG 2157

RESULT 15
 ABS65377
 ID
 ABS65377 standard; DNA; 2158 BP.
 XX

AC ABS65377;
 XX 15-NOV-2002 (first entry)
 DT Clostridium difficile strain 170324 slpA DNA sequence.
 XX Vaccine; Clostridium difficile; immunogenic; immunoreactivity;
 XX C. difficile infection; immune response; C. difficile associated disease;
 KW passive immunotherapy; passive vaccination; interleukin; adjuvant;
 KW C. difficile vaccine; antibacterial; slpA; ds.
 XX Clostridium difficile.
 OS WO200262379-A2.
 PN 15-AUG-2002.
 XX 11-FEB-2002; 2002WO-IE000017.
 XX 09-FEB-2001; 2001IE-00000137.
 PR (QUBE-) QUEEN ELIZABETH COLLEGE DUBLIN.
 XX Doyle R, Kelleher D, Windle HJ, Walsh JB, Deirdre NE;
 PI WPI; 2002-643389/69.
 DR Novel vaccine useful for treatment/prophylaxis of Clostridium difficile
 PT associated diseases, comprises Clostridium difficile genes or
 PT peptide/polypeptides or its derivative, fragment, mutant or variant.
 XX Claim 13; Page 71-72; 85pp; English.
 XX The present invention relates to a new vaccine for treatment/prophylaxis
 CC of Clostridium difficile associated disease. The vaccine of the invention
 CC comprises C. difficile gene or peptide/polypeptide or derivative,
 CC fragment, mutant or variant of the peptide/polypeptide which is
 CC immunogenic in humans or to which immunoreactivity is detected in
 CC individuals who have recovered from C. difficile infection. The molecules
 CC of the invention are used in prophylaxis or treatment of C. difficile
 CC associated disease, where the medicament is a vaccine. The invention can
 CC be used by administering the vaccine preparation to a host to raise an
 CC immune response. The invention is further useful in the preparation of a
 CC medicament for treatment or prophylaxis of C. difficile infection or C.
 CC difficile associated disease. The invention is also useful in passive
 CC immunotherapy for established C. difficile infection, for the eradication
 CC of C. difficile associated disease, or for passive vaccination of an
 CC individual with C. difficile infection. Interleukin is useful as an
 CC adjuvant in C. difficile vaccine. The present nucleic acid sequence
 CC represents a Clostridium difficile slpA DNA sequence of the invention
 XX Sequence 2158 BP; 843 A; 246 C; 435 G; 634 T; 0 U; 0 Other;

Query Match 25.8%; Score 471.8; DB 6; Length 2158;
 Best Local Similarity 61.5%; Pred No. 3.3e-77;
 Matches 951; Conservative 0; Mismatches 467; Indels 129; Gaps 7;
 QY 296 CTACAGCTAAATGGAATGAAGATTATGTAAGACAACTTTAAAAATTTAGATGCGAGG 355
 Db 728 CAACACCTAGTGTCTAGCTGTAAGCTGGTTTGTAACTAAAGATGATCTAGTTAGCAA 787
 QY 356 ATATGCTTATATAGATTTAACCTTATATATGCTTAAACTGTTGAAATTAAGTAGTAG 415
 Db 788 AATCAGGTACTATAAATGTAAGATTAAATGCAAAAAGAAATCAATTTGATATAGATG 847
 QY 416 CAGCTAGTGAATAAACAAGTGTGTATCTAGTGTGCGAAAAATAGTGCAAAAGATATAG 475
 Db 848 CAGCTCATATACATCAGCTG-----AAATTTAGCTTAAAGATATGTTTGTATCCAG 901
 QY 476 CTGAAAAATATGTTGTTTGAAGCAAAAGACTTTAGAAAAATGCATAAAAACTATAAATGCCT 535
 Db 902 ATGAAATTTCTGAACATATAAGGCAATAGTACGATTACAAATATGATGTATAGATCTTA 961

Qy 536 CAGATTTCAGTAAACTGATAGTTACTATCAAGTAGTCTTCTTATCCAAAGGAAAGAGAT 595
 Db 962 ACTTAGTTCAGTTAGTTAATGGAATAATCAAGTGATTTTATCCAGAGGTAAAGAT 1021
 Qy 596 TACAAGGTTTCTCAACTATATAGAGCTACAAATTAATGAAGAACTGCATATGGTAATA 655
 Db 1022 TAGAACTAATCAGCA-----AATGATACAATAGCTAGTCAAGATA 1063
 Qy 656 CACAGTAAATTAACCTTAATACTACTAGTAAGAGTAATTTAAAGACTGCAGTAGAAG 715
 Db 1064 CACAGCTAAGTAGTTATAAAGCTAAATAATTAAGATTTTAAAGATTTATGTAGATG 1123
 Qy 716 AGTTACAAAATTTGAATGCTAGTTATTTCTAATACTACAACTTTAGCTGGTGATGACAGAA 775
 Db 1124 ATTTAAACATATATATATCTTATTTCAATGTTGTACAGTAGCAGGAGATAGAA 1183
 Qy 776 TACAAACAGCTATAGAGATAAGTAAGTAAGAAATTTATCAATATATGATGGGAGAAATCAGATC 835
 Db 1184 TAGAACTGCTATAGAAATTAAGTAGTAATATTAATAATCTGAT-----GATA 1231
 Qy 836 ATTCAGCTGATTTAAAGAGAAATTTAAATAATGTTGTATTTAGTAGTGGCAATGCACTAG 895
 Db 1232 AAAATGCAATAACTGATAAGCAAGTAAATGATATAGTATTTAGTTGGATCTACATCTATAG 1291
 Qy 896 TAGATGATTTAGTTGGGCTCTTTAGCAGCAGAAAGATGCTCCACTATTTAATACTT 955
 Db 1292 TTGATGCTCTGTTGTCATCACCATTAGCTTCAGAAATAACAGCTCCATTTATTTAATTT 1351
 Qy 956 CAAAAGATAAATTAGATTCCTGAGTAAATCTGAAATTAAGAGAGTTTTAGACTTAAAAA 1015
 Db 1352 CAAAAGATAAATTAGATTCATCAGTAAATCTGAAATTAAGAGAGTTATGAACTTAAAGA 1411
 Qy 1016 CTTCAACAGAGTA---ACAGGAAAAACAGTTTTATATAGCTGGTGGAGTTAATAGTAT 1072
 Db 1412 GTGACACTGGTATAAATACTTCTAAAAAGATTTATTTAGCTGGTGGAGTTAATTTCTATAT 1471
 Qy 1073 CTAAGAGATTGTAAACAGAAATTAAGATCAATGGGATTAAGATTTGAAAGATTTCTCAGGTG 1132
 Db 1472 CTAAGATGTAGAAATGAAATTTGAAACAAATGGGCTTTAAAGTTACTAGATTTACAGGAG 1531
 Qy 1133 ATGATAGATATGAACTTCTTTAAAAATAGCAGGTGAAATPAGGCTTAGATAAATGATAAGG 1192
 Db 1532 AAGACAGATACGAACTCTTTAGCAATAGCTGATGAAATAGGTCTTTGATAATGATAAAG 1591
 Qy 1193 CTTATGATGTTGGTGGAAACAGGATTAGCAGATGCCATGAGTATAGCTTCAGTTGCTTCTA 1252
 Db 1592 CATTTGTAGTTGGTGGTACTGGATTAGCAGATGCTATGATATAGCTCCAGTTGCTTCT- 1650
 Qy 1253 CTAATTTAGATGGTAAATGGTGTGTTAGATAGAACAAATGGACATGCTACTCCAATAGTTG 1312
 Db 1651 -----CAACTTAAAGATGGAGATGCTACTCCAATAGTAG 1684
 Qy 1313 TTGTAGATGGAAGAGCTGATAAAATATCTGATGACTTTAGATAGTTTCTTAGGAAGCGCTG 1372
 Db 1685 TTGTAGATGGAAGAGCAAGAAATAGTGTGATGCTAGAGTTTCTTAGGAACCTTCTG 1744
 Qy 1373 ATGTAGATATATAGTGGATTTGCAAGTGTATCTGAAAGAGATGGAAAGAGCTATATCAG 1432
 Db 1745 ATGTTGATATAATAGTGGAAAAAATAGCGTATCTAAAGAGATTTGAAGAGTCAATAGATA 1804
 Qy 1433 ATGCTACTGGTAAAGCGGTTACAGAGTTTAAAGCGGACGATAGACAGACACTAACTCTG 1492
 Db 1805 GTGCAACTGGAAAAAATCTCCAGATAGAAATAGTGGAGATGATAGACAAGCAACTAATGCTG 1864
 Qy 1493 AAGTTATAAAAAATATTTATCTAATGATGATACTGAAATAGCTTAAAGCTGCAGTTTTAGATA 1552
 Db 1865 AAGTTTAAAGAGAGATGATTTTC----- 1890
 Qy 1553 AAGATTCAGGTGCTTCAAGTAGTATGATGAGAGATATTTAATTTCTATGCTAGCTAAAGATG 1612
 Db 1891 -----ACAGATGGTGAAGTTGTGAATTTACTTTGTTGCAAAAAGATG 1930
 Qy 1613 GATCTACAAAAGAGATCAATTTAGTTGATGCAATTAGCAGTAGGAGCTGTTGCTGG----- 1667

Db 1931 GTTCTACTAAAGAGATCAATTTAGTAGATGCCCTTAGCAGCAGCAACAATAGCAGGTAGAT 1990
 Qy 1668 ----ATATAAACTTGCCTCCAGTTGTATTAGCTACTGATTTCTTTATCTTCTGATCAATCGG 1723
 Db 1991 TTAAGGAGTCTCCAGCTCCATCATCATCTAGCTACTGATCTTTATCTTCTGACCAAAATG 2050
 Qy 1724 TTGCTATTAAGCAAAAGTTGTAGGAGAAAAATATTTCTAAAGATTTAAACAAGTTGGTCAAG 1783
 Db 2051 TAGCTGTAAGTAAAGCAGTTCCCTAAAGATGGTGGAACTAACTTAGTTCAAGTAGGTAAAG 2110
 Qy 1784 GAATAGCTTAATTCAGTTATTAACAACAAATGAAAGATTTATTAGATATG 1830
 Db 2111 GTATAGCTTCTTTCAGTTTATAACAAATGAAAGATTTATTAGATATG 2157

Search completed: October 6, 2005, 06:23:53
 Job time : 1048 secs

This Page Blank (uspto)